NASA

164000-18-P

Tenth Type II Quarterly Status
and Technical Progress Report

STUDY OF SPECTRAL/RADIOMETRIC CHARACTERISTICS OF THE THEMATIC MAPPER FOR LAND USE APPLICATIONS

21 December 1984 — 20 March 1985

WILLIAM A. MALILA
MICHAEL D. METZLER

MAY 1985

Contract NAS5-27346
NASA Goddard Space Flight Center
Greenbelt Road
Greenbelt, MD 20771

ENVIRONMENTAL
RESEARCH INSTITUTE OF MICHIGAN
BOX 8618 • ANN ARBOR • MICHIGAN 48107

NOTICES

Sponsorship. The work reported herein was conducted by the Environmental Research Institute of Michigan under Contract NAS9-16538 for the National Aeronautics and Space Administration, Johnson Space Center, Houston, Texas 77058. Victor S. Whitehead was Technical Monitor for NASA. Contracts and grants to the Institute for the support of sponsored research are administered through the Office of Contracts Administration.

Disclaimers. This memorandum was prepared as an account of Government sponsored work. Neither the United States, nor the National Aeronautics and Space Administration (NASA), nor any person acting on behalf of NASA:

(A) Makes any warranty expressed or implied, with respect to the accuracy, completeness, or usefulness of the information, apparatus, method, or process disclosed in this memorandum may not infringe privately owned rights; or

(B) Assumes any liabilities with respect to the use of, or for damages resulting from the use of any information, apparatus, method, or process disclosed in this memorandum.

As used above, "person acting on behalf of NASA" includes any employee or contractor of NASA, or employee of such contractor, to the extent that such employee or contractor of NASA or employee of such contractor prepares, disseminates, or provides access to any information pursuant to his employment or contract with NASA, or his employment with such contractor.

Availability Notice. Requests for copies of this memorandum should be referred to:

National Aeronautics and Space Administration
Scientific and Technical Information Facility
P.O. Box 33
College Park, Maryland 20740

Final Disposition. After this document has served its purpose, it may be destroyed. Please do not return it to the Environmental Research Institute of Michigan.

TECHNICAL REPORT STANDARD TITLE PAGE

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
164000-18-P		

4. Title and Subtitle	5. Report Date
Study of Spectral/Radiometric Characteristics of the Thematic Mapper for Land Use Applications	May 1985
	6. Performing Organization Code
7. Author(s)	8. Performing Organization Report No.
William A. Malila and Michael D. Metzler	164000-18-P
9. Performing Organization Name and Address	10. Work Unit No.
Environmental Research Institute of Michigan P.O. Box 8618 Ann Arbor, MI 48107	11. Contract or Grant No. NAS5-27346
	13. Type of Report and Period Covered Quarterly Status and Technical Progress 21 Dec 1984 - 20 Mar 1985
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Goddard Space Flight Center Greenbelt, MD 20771	14. Sponsoring Agency Code

15. Supplementary Notes

Mr. Harold Oseroff (Code 902) is serving as Technical Officer and Mssrs. Brian Markham and James Irons (Code 923) are serving as Science Representatives for NASA/GSFC.

16. Abstract

Progress during ERIM's tenth quarter of effort under the Landsat-4 and 5 Image Date Quality Assessment program for the Thematic Mapper is described.

Coincident Landsat-4 and 5 fully corrected (CCT-PT) TM data were analyzed in more detail and revised band-by-band relationships between the two sensors were derived. An analysis technique employing the matching of cumulative distributions was developed and used and is believed to offer advantages over the histogram matching procedure currently used to produce Landsat data. Multiplicative factors ranging from 0.987 to 1.145 and offsets ranging from -2.7 to -6.2 video quantum levels were required to cause Landsat-5 data to match Landsat-4 data values. Evidence of low-level clipping was found in TM Bands 5 and 7 of Landsat-5 but not Landsat-4.

Analysis of the information content of Landsat TM and MSS data was continued. Components of information loss were identified and quantified and the effects of coarsened quantization were explored.

17. Key Words	18. Distribution Statement
Radiometric Calibration Landsat-4, Landsat-5 Thematic Mapper Noise	Initial distribution is listed at the end of this document.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
UNCLASSIFIED	UNCLASSIFIED	69 + v	

Report No. 164000-18-P

Tenth
Type II Quarterly Status
and Technical Progress Report
21 December 1984 – 20 March 1985

on

Study of Spectral/Radiometric Characteristics of the Thematic Mapper for Land Use Applications

under

Contract NAS5-27346

with

NASA Goddard Space Flight Center
Greenbelt Road
Greenbelt, Maryland 20771

Submitted by

Environmental Research Institute of Michigan
P.O. Box 8618
Ann Arbor, Michigan 48107

Prepared by: William A. Malila
William A. Malila
Principal Investigator

Michael D. Metzler
Co-Investigator

Approved by: _____
Robert Horvath
Manager, Information Processing
Department

May 1985

Table of Contents

Tenth Quarterly Report

STUDY OF SPECTRAL/RADIOMETRIC CHARACTERISTICS OF THE THEMATIC MAPPER FOR LAND USE APPLICATIONS

1. OBJECTIVE

The objective of this investigation is to quantify the performance of the TM as manifested by the quality of its image data in order to suggest improvements in data production and to assess the effects of the data quality on its utility for land resources applications. Three categories of this analysis are: a) radiometric effects, b) spatial effects, and c) geometric effects, with emphasis on radiometric effects.

2. TASKS

Four tasks have been established to address the above objective. The first three are to study radiometric performance, spatial performance, and geometric performance, respectively, while the fourth is to study spectral characteristics. In keeping with the identified objective, the radiometric performance study is our major task.

3. STATUS AND TECHNICAL PROGRESS

During this tenth quarterly reporting period, analysis was continued of coincident fully-corrected Landsat-4 and Landsat-5 TM data. Also, additional analyses of the information content of Landsat data were conducted.

3.1 PROBLEMS

None.

3.2 ACCOMPLISHMENTS

Accomplishments in two technical areas are summarized below and described in detail in Appendices A and B.

3.2.1 TM Landsat 4 vs Landsat 5 Radiometric Comparison

During the previous quarter[1], preliminary relationships were established between fully-corrected (CCT-PT) data from coincident frames of Thematic Mapper data from Landsats 4 and 5. Revised relationships were developed this quarter, after further analysis of the data set. Two techniques were employed to compare the coincident data. The first technique involved spatial registration of a 3.6×10^6-pixel subimage (1980 lines by 1800 pixels) of the Landsat-5 frame to the Landsat-4 frame, spatial averaging to reduce any misregristration effects and to reduce data volume, and band-by-band linear regressions of data from the two averaged images. The second technique employed a novel histogram matching procedure. This procedure matches the cumulative distribution functions of paired histograms, rather than equalizing only the means and standard deviations as is done by the TM ground processing system in its histogram equalization procedure. The cumulative-distribution technique is believed to have some advantages in making better use of the full range and distribution of data values in the histogram. Additional details of the analysis are presented in Appendix A, and the results are summarized below.

Since both the Landsat-4 and Landsat-5 scenes were processed through TIPS (*T*hematic *M*apper *I*mage *P*rocessing *S*ystem), it was expected that radiometrically corrected products would have essentially identical corrected signal values for the same scene viewed at the same time. However, substantial differences were found and clipping of the Landsat-5 data values was obvious in both Bands 5 and 7 at the low radiance end of the dynamic range. The Band 7 low-level clipping is apparent from a histogram of signal-level frequency for Band 7 for both Landsat 4 and Landsat 5 (see Figure 1). The pixels with amplitude values zero to six in the Landsat-4 scene are all mapped to value zero in the Landsat-5 scene. Although the offset was nearly as large for Band 5 (see Figure 2), fewer data values actually were clipped (0.3 percent of the scene versus 4.2 percent in Band 7).

As noted above, band-by-band comparisons were carried out using two different techniques. When clipping was not present, the techniques produced essentially the same results. Where clipping was present, regression of matched areas led to smaller offsets and larger multiplicative factors, a result deemed erroneous after inspecting the histograms. For this reason, coefficients from the cumulative-distribution matching approach are presented here. Table 1 presents the coefficients to convert Landsat-4 TM signal levels to Landsat-5 TM equivalent values and Table 2 contains the coefficients to convert Landsat-5 signals to Landsat-4

values. (Coefficients to match radiances are included in Appendix A.) It should be noted that while this was a simultaneously collected data set, and therefore nearly ideal for this type of analysis, the coefficients presented are valid only if ground processing parameters are not changed. It should also be noted that data which have been clipped as in Landsat-5 Bands 5 and 7 can not be retrieved — all the zeroes in Landsat-5 Band 7 data will be converted to sixes in Landsat-4 Band 7 whereas Landsat-4 Band 7 would have recorded the same pixels with signal levels ranging from zero to six.

TABLE 1. Coefficients for Converting Landsat−4 Values to Landsat−5 Values
(Scenes 4−0608−15463 and 5−0014−15460, 15 March 1984)

Landsat-5 TM = A*(Landsat-4 TM) + B

Band	A	B	S.E.	R^2	Range of Data Values Landsat-4	Landsat-5
1	1.0438	−3.538	0.151	0.99943	73-109	73-111
2	1.1200	−2.719	0.134	0.99922	26-52	26-56
3	0.9869	−3.678	0.142	0.99975	26-77	22-72
4	1.0030	−4.627	0.078	0.99995	11-92	7-88
5	1.1452	−7.330	0.106	0.99999	6-154	0-169
6	1.0040	−0.711	0.119	0.99956	114-148	113-148
7	1.0923	−6.244	0.054	0.99999	3-86	0-88

Note: If the value computed for Landsat-5 is <0, substitute 0.
If it is >255, substitute 255.

TABLE 2. Coefficients for Converting Landsat−5 Values to Landsat−4 Values
(Scenes 4−0608−15463 and 5−0014−15460, 15 March 1984)

Landsat-4 TM = A*(Landsat-5 TM) + B

Band	A	B	S.E.	R^2	Range of Data Values Landsat-4	Landsat-5
1	0.9580	3.390	0.145	0.99943	73-109	73-111
2	0.8928	2.427	0.120	0.99922	26-52	26-56
3	1.0132	3.726	0.144	0.99975	26-77	22-72
4	0.9970	4.614	0.078	0.99995	11-92	7-88
5	0.8732	6.401	0.093	0.99999	6-154	0-169
6	0.9960	0.714	0.118	0.99956	114-148	113-148
7	0.9155	5.717	0.049	0.99999	3-86	0-88

Note: If the value computed for Landsat-4 is >255, substitute 255.

3.2.2 Information Content of Landsat TM and MSS Data

Previously, a relative-entropy relationship was developed to measure the information content of multispectral scanner data[2]. The maximum value would result if each observation were in a unique spectral cell. Information is lost when the observations cluster or concentrate into a smaller subset of spectral cells. During this reporting period, the nature of this information loss was examined and two components were identified. The first, called cell loss, is due to the reduction in number of cells below the number of observations. The remainder and second component, called uniformity loss, occurs when the duplicate observations are not uniformly distributed among the occupied cells. These components were evaluated for several areas extracted from a Landsat frame, ranging in size from 10^5 to 10^6 TM pixels. Detailed results are presented in the paper in Appendix B. Information losses for TM data were notably smaller than for corresponding MSS data, and the relative entropy values for TM data were larger.

A second study also was conducted, to explore the information content under coarsened quantization of the signal amplitudes. This is related to the effects of noise which, by adding variance to signals, could increase the number of spectral cells occupied and create an apparent information content greater than that present in ideal, noiseless signals. Again, details of this analysis are presented in the paper in Appendix B. The TM data sets maintained higher relative entropy values when signal quantization was degraded to seven and six bits/band, but the MSS and TM values were about equal when both were degraded to five bits/band.

3.3 SIGNIFICANT RESULTS

Substantial differences were observed in the calibrated amplitudes of fully corrected Landsat-4 and -5 TM data from a coincident scene. Multiplicative factors ranging from 0.987 to 1.145 were required to cause Landsat-5 data values to match Landsat-4 values, along with offsets ranging from -2.7 to -6.2 video quantum levels. Low-level clipping was found in Bands 5 and 7 of the radiometrically corrected Landsat-5 data.

An improved histogram matching algorithm was developed and applied to produce the revised relationships between amplitude values from Landsats 4 and 5.

3.4 PUBLICATIONS AND PRESENTATIONS

W. Malila and M. Metzler attended and made presentations at the Landsat-5 TM Investigators Workshop, January 9-10, 1985 at the NASA Goddard Space Flight Center.

Two papers submitted and accepted for publication in the issue of Photogrammetric Engineering and Remote Sensing which will feature papers from the final Landsat Image Date Quality Analysis (LIDQA) symposium to be held in Indianapolis, Indiana, in September, 1985. A third paper will also be presented at the symposium. Preliminary copies of the two papers are included in Appendices A

and B.

3.5 RECOMMENDATIONS

It is recommended that NASA and/or NOAA investigate the reasons for the differing values of coincident data sets after radiometric correction and make appropriate adjustments.

3.6 FUNDS EXPENDED

A total of approximately $17,000 was expended during the three months December 1984 through February 1985. The cumulative spending through February represents approximately 86% of the contract award. Expenditures during the period 1-20 March 1985 are not included in this percentage value.

3.7 DATA RECEIPTS

Raw data tapes (unity RLUT CCT-AT) and calibration data tapes (CALDUMP) were received during this quarter for the following scenes:

San Francisco	P44/R34	5-0062-18131
White Sands	P33/R37	5-0129-17075

Radiometrically corrected data (CCT-AT) were received for six scenes:

NE Alabama	P20/R36	5-0014-15454
Iowa	P28/R30	5-0046-16324
Harrisburg	P111/R212	5-0052-02182
San Francisco	P44/R34	5-0062-18131
San Francisco	P44/R34	5-0126-18143
Iowa	P27/R31	5-0151-16290

Fully corrected data (CCT-PT) were received for two scenes:

Iowa	P28/R30	5-0046-16324
San Francisco	P44/R34	5-0062-18131

REFERENCES

1. Malila, W.A., and M.D. Metzler, *Ninth Quarterly Progress Report for Study of Spectral/Radiometric Characteristics of the Thematic Mapper for Land Use Applications*, ERIM Report No. 164000-14-P, January 1985.

2. Malila, W.A., *Sixth Quarterly Progress Report for Study on Spectral/Radiometric Characteristics of the Thematic Mapper for Land Use Applications*, ERIM Report No. 164000-10-P, April 1984.

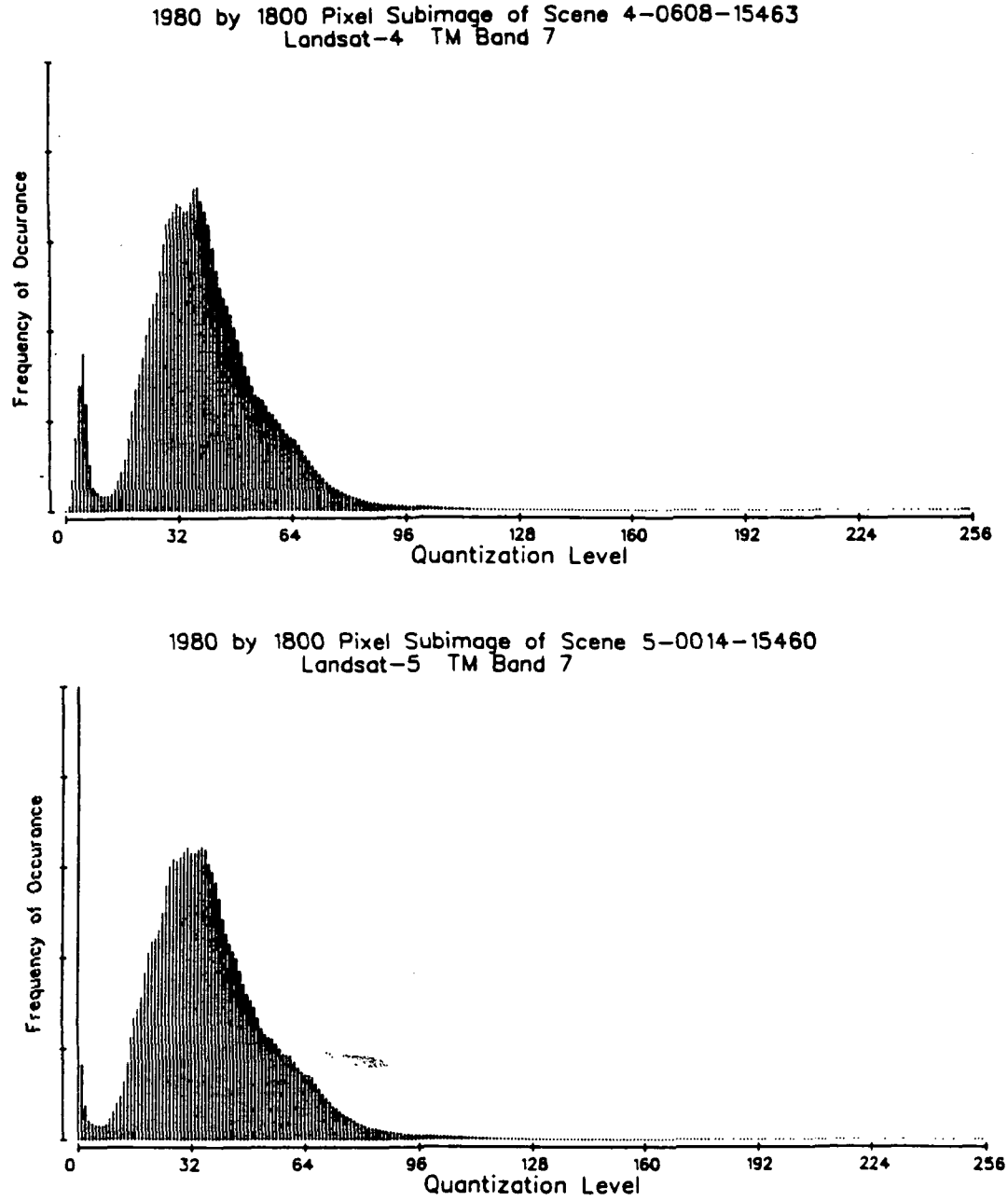


Figure 1. TM Band-7 Histograms for Region Viewed Simultaneously by Landsats 4 and 5.

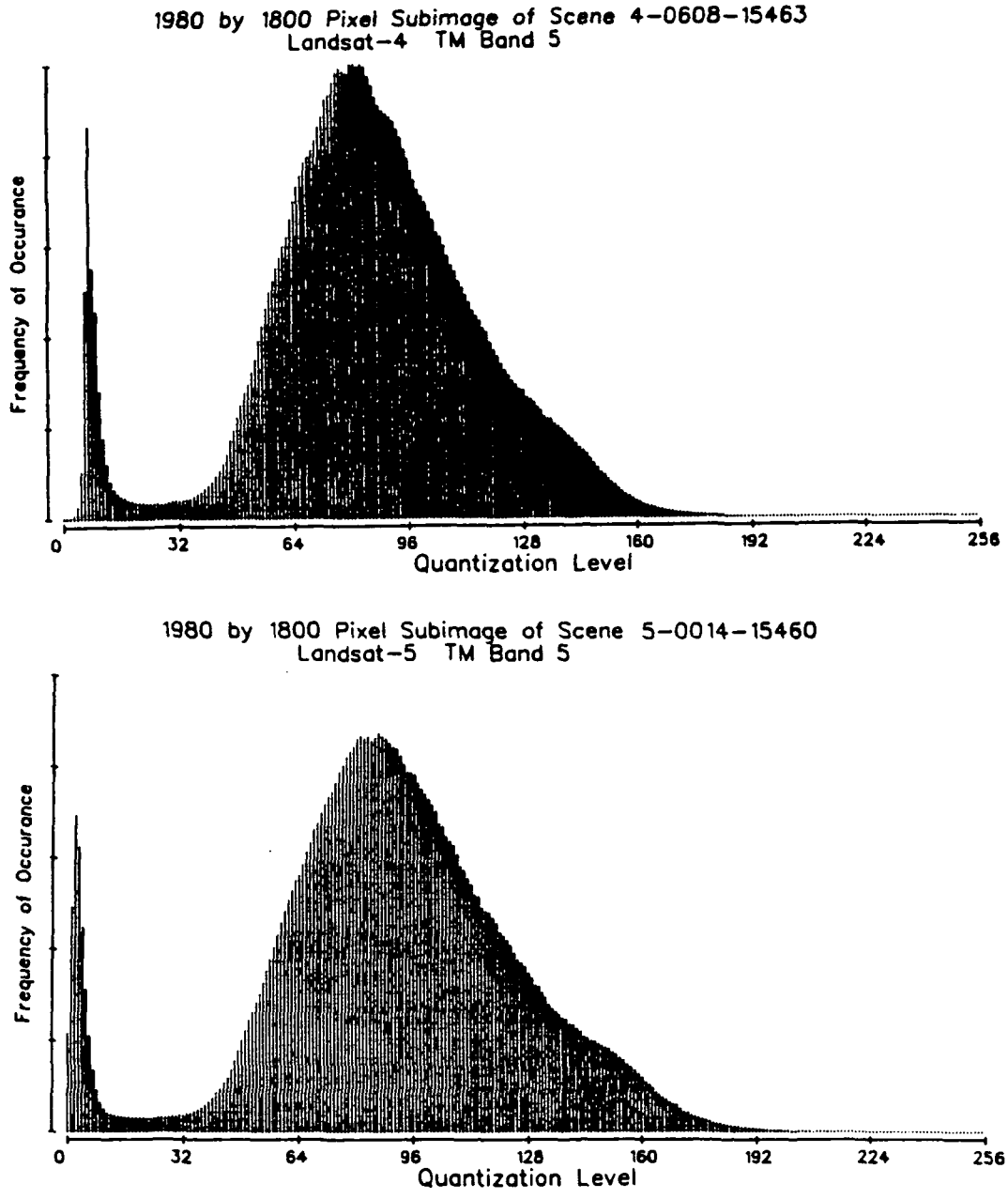


Figure 2. TM Band-5 Histograms for Region Viewed Simultaneously by Landsats 4 and 5.

APPENDIX A

Michael D. Metzler
William A. Malila
Environmental Research Institute of Michigan
P.O. Box 8618 Ann Arbor, MI 48107

Characterization and Comparison of Landsat-4 and -5 Thematic Mapper Data

Engineering analyses of Thematic Mapper image data from Landsats 4 and 5 were performed with an emphasis on radiometric performance.

ABSTRACT

Analyses of the characteristics of Landsat Thematic Mapper (TM) image data are described and results are summarized. Emphasis is placed on radiometric characterization, development of response models, and on comparisons between data from Landsats 4 and 5. In general, the data quality was excellent; however, some anomalies were found. Three main topics are (a) systematic within-scan-line signal droop/rise, (b) random scan-correlated level shifts, and (c) radiometric (signal amplitude) relationships between Landsat 4 and 5. The systematic droop/rise effect was found in data from both Landsats 4 and 5. Daytime signals droop across the scan line while nighttime signals in the reflective bands rise across the scan line. The magnitude of the droop/rise appears to be a function of the signal magnitude and average value of the signal throughout a scan cycle. Scan-correlated level-shift noise also was observed in data from both sensors, but with different patterns. Low-amplitude, low-frequency coherent noise effects also were measured. The analysis of simultaneously acquired Landsat-4 and -5 TM data permitted a direct empirical comparison of the relative radiometric responses of their respective spectral bands. Relationships between their respective signal values were developed, and sensor dynamic range considerations are discussed. It was determined that multiplicative factors ranging from 0.987 to 1.145 were required to convert the signal counts from Landsat-4 TM spectral bands to corresponding Landsat-5 equivalent signals. Radiance values exhibited corresponding differences, pointing to residual errors in radiometric calibration. Low-level clipping was evident in the radiometrically corrected Landsat-5 Band 5 and 7 data. The temperature range covered by the full 8-bit data range of TIPS-processed TM Band 6 data was found to be approximately 200°K to 340°K, not 260°K to 320°K as specified.

This research was sponsored by the U.S. National Aeronautics and Space Administration, Goddard Space Flight Center, Greenbelt, MD, under Contract NAS5-27346.

1. INTRODUCTION

Since the launch of Landsat 4 in July 1982, numerous studies of the quality of Thematic Mapper image data have been performed under the auspices of NASA's Landsat Image Data Quality Analysis (LIDQA) program. As part of this program, we have performed engineering analyses of Thematic Mapper image data with our efforts concentrated on radiometric characterization of the sensor. In general, we have found the data quality to be excellent. However, anomalies do exist in the data from both Landsat-4 and Landsat-5 TM. The analyses of Landsat-4 TM image data were previously described in detail by the authors [Metzler and Malila 1983, Malila et.al. 1984, and Metzler and Malila 1985], and are summarized below. This paper concentrates on recent analyses of Landsat-5 TM data, and comparisons of the radiometry of the two sensors. Specifically, topics covered are: (a) within-line 'droop', a phenomenon whereby the signal levels of the sensor change systematically during the active scan, (b) scan-correlated level shifts, an effect which raises or lowers the signal level of all pixels in a scan line or set of scan lines, and (c) comparison of Landsat-4 and Landsat-5 radiometric corrections. Other analyses of TM data anomalies may be found elsewhere in this issue (e.g., Kieffer et.al., 1985).

Within−Line Droop. Earlier examination of Landsat-4 TM 'average' scan lines indicated significant differences in the signals returned from the Western edge of a scene compared to those observed at the Eastern edge of the same scene. This effect was most apparent in the shortest wavelength spectral band (Band 1), and was observed in all the spectral bands to some extent. A combination of bidirectional reflectance, atmospheric, and shadowing effects, as well as sun-view angle geometry can explain the effect observed. More careful examination of the 'average' scan data, however, revealed a confounding effect due to different sensor response characteristics related to the direction of scan in Bands 1-4. The scan-direction difference took the form of a 'droop' in signal with time during active scan, which appeared as a signal decrease with increasing pixel number for forward scans, and a signal increase with increasing pixel number for reverse scans. This effect was found in nighttime reflective-band data as well, but taking the form of a signal 'rise' with time instead of a 'droop'.

Scan−Correlated Level Shifts. In Landsat-4 TM data, an effect was analyzed which changed the signal of all samples within a scan-line or group of scan-lines by up to 2.0 video quantum levels (DN). The changes were aperiodic, occurring at random intervals with the level shifting during mirror turn-around time. All affected detectors shifted levels at the same time, with the level shifts following one of two patterns (most detectors exhibited both patterns, but one was dominant). One pattern was exemplified by Band 1 Detector 4 with a peak-to-peak amplitude of 2.0 DN, the other by Band 7 Detector 7. These two patterns were labeled 'form #1' and 'form #2', respectively (later labeled 'Type 4-1' and 'Type 4-7', respectively, by Barker [1984]).

2. METHODS

All analyses to characterize the radiometry of the sensors were performed on digital computer-compatible-tape (CCT) data. Several types of CCT data were used, representing various stages of ground processing as well as calibration data. The analyses described in this paper generally were performed on full-frame TM image data, both to characterize full-frame effects and to take advantage of the large data volume (approximately 37 million pixels per band per frame) to improve the quality of the statistics generated.

Two primary methods were used to average the full-frame image data. In one case, to examine scan-angle effects, 'average' scan lines were computed by averaging the columns of pixels down the entire frame. To analyze scan-direction effects, these 'average' scan lines were stratified by scan direction, with the forward and reverse scan data being treated separately. The other type of averaging involved computing down-track profiles by averaging the rows of pixels across the entire frame, thereby computing an average signal value for each scan line. Each of these analyses was performed seperately for each band and each detector of the sensors.

Earlier investigations by the authors [Malila et.al. 1984] demonstrated the value of using reflective-band (Bands 1-5 and 7) TM data collected at night for analysis of sensor data anomalies. The sensor sensitivities are such that no scene radiance is recorded, so any variations in the data are due to sensor noise effects. We again made extensive use of nighttime data in the analyses described herein, processing the data using the techniques described above.

Two techniques were employed to compare the radiometry of the Thematic Mappers on Landsats 4 and 5, using a special data set which was collected simultaneously with both sensors. The first technique involved selecting a sub-image of 3,564,000 pixels (1980 lines by 1800 pixels) from the Landsat-5 image and spatially registering it to the Landsat-4 TM image data. This registration was performed to sub-pixel accuracy using 50 control points and nearest neighbor resampling. The data in each sub-image were averaged using five-by-five-pixel cells to reduce any misregistration effects and to reduce the data volume, while still retaining the data diversity. Linear regressions were performed with data from the two averaged images. Multiplicative and additive factors were computed for each band which can be used to relate the signals from one sensor to those expected from the other sensor for the same input radiance.

For the second comparison, histograms were computed for the sub-images described above, and a histogram matching technique was employed to compute multiplicative and additive coefficients for relating data from one sensor to that of the other. Unlike the histogram matching procedure used in TM ground processing which equalizes means and standard deviations [Barker et.al. 1985], a procedure based on matching the cumulative distribution functions of the two data sets was employed. This technique is believed to have some advantages in making better use of the full range and distribution of data values in the histogram. Additionally, the

histogram matching approach has much less stringent registration requirements than pixel or region matching approaches. For a region with a very small perimeter/area ratio, effects of slight misregistration would be minimal. The cumulative distribution function gives percentage of observations having signal values less than or equal to the designated signal value. Interpolations were made to obtain the signal values corresponding to integer percentage values. Excluding end points, a regression of the corresponding percentile signal levels from the two sensors provided the desired correction coefficients.

3. RESULTS

Within−Line Droop. The single nighttime Landsat-5 TM scene (ID 5-0052-02182, Harrisburg) available to us was used to quantify the within-scan 'droop/rise' effect in Landsat-5 TM data. The 'average' nighttime scan lines for Band 4 of Landsat-5 TM are illustrated in Figure 1, along with data from Landsat-4 TM for comparison. Both forward and reverse scans are shown. The *y*-axes all have the same scale, i.e. 0.1 DN full scale, to facilitate comparison between sensors. Note that for reverse scans, pixel Position 6000 is sampled prior to pixel Position 1. Therefore, the effect is seen to be a signal 'rise' with increasing time for both forward and reverse scans. In general, the within-scan 'rise' has the same magnitude and time constant for the same band in each sensor. Magnitudes are greater in daytime data and the signals 'droop' with increasing time, as will be discussed later.

Band 1 displays the greatest effect, with the mean reverse-scan signal increasing approximately 0.1 DN during the active scan. A simplified exponential decay model was fitted to the data for each of Bands 1-4. For these bands, the time constant (time for magnitude of effect to decay to *1/e* of original value) which produced the best fit ranged from 900 to 1100 pixel sample times, (approximately 9-10 milliseconds) for both Landsat-4 and -5 TM.

The mathematical model used is expressed by the equation:

$$S(p) = S_0 + Be^{-p/T}$$

Where:

S(p) = signal returned by sensor for pixel 'p'
S_0 = signal for 'p' equal to infinity
B = magnitude of total 'droop/rise'
T = time (pixels) required for signal to change by 63% of 'B'
p = pixel number, with count starting with first image pixel (West-most for forward scans, East-most for reverse daytime scans)

Since the magnitude and time constant of the nighttime within-scan 'rise' are essentially identical for Landsats 4 and 5, we would expect the daytime 'droop' effects to be similar also. During daytime data acquisition when signal levels were much higher, we observed in Landsat-4 TM data a corresponding increase in the

magnitude of the 'droop' effect. In a daytime Band 1 scene (ID 4-0049-16262) which had a scene mean of 61.9 DN, the magnitude of the 'droop' was observed to be approximately minus 1.5 DN, with a time constant equivalent to approximately 900 pixels. At night, the magnitude of the 'rise' was <0.15 DN, still with a time constant of 900 pixels for Band 1. The mean scene level at night was 2.3 DN. Although qualitatively the daytime Landsat-5 effect appears similar to the daytime Landsat-4 effect, quantification of this effect in daytime Landsat-5 TM data awaits analysis of an appropriate scene in which variations in scene radiance have a relatively uniform spatial distribution.

The 'droop/rise' effect was analyzed further to establish a hypothesis for its cause and a model for its description and potential use in correction. While the magnitude of the effect does not appear to be strictly proportional to the scene mean, it does appear as if the 'droop' or 'rise' is a drift toward the 'scan-cycle mean' signal of the scene which also includes the signal values produced during shutter obscuration, calibration pulse, and DC restoration. This 'scan-cycle mean' would be lower than the scene mean during the daytime due to the addition of the data acquired during shutter obscuration, and would be greater than the scene mean during nighttime data acquisition, where the scene itself is effectively a continuation of the shutter obscuration, and the calibration pulses drive the 'scan-cycle mean' to a level slightly higher than the scene mean (see Figure 2). The hypothesis is that a-c coupling exists between the detector output and the analog-to-digital converter, producing a signal decay proportional to the departure from the scan-cycle mean.

To test this hypothesis, a scene of nighttime Landsat-4 TM data (Scene 4-0037-02243) was segmented based on the calibration lamp state observed by the sensors prior to each scan. As illustrated in Figure 3, the internal calibration lamps sequence through the eight possible states, remaining in each state for approximately 40 scans (20 forward/reverse scan cycles). During nighttime reflective-band data collection, the signal pulses resulting from viewing these calibration lamps at the end of each scan are the only signals available to shift the 'scan-cycle mean' from the scene mean. All scans with lamp state 000 (no lamps on) were grouped into one sub-image, and seven other sub-images were created for the other seven lamp states (001, 010, 011, 100, 101, 110, and 111, where each binary digit represents the state of one of the three calibration lamps). 'Average' scan-lines were computed for each of these sub-images, then smoothed and displayed as plots of mean signal level vs pixel position (see Figures 4a-4h). Qualitatively one can see that the effect is greatest when the calibration pulse adds the most to the 'scan-cycle mean' (state 111, all lamps on), and is non-existent in the case of no calibration pulse ('scan-cycle mean' equal to scene mean).

Quantitative support for the hypothesis of drift toward a 'scan-cycle mean' was derived from data on the calibration tape (CCT-ADDS) associated with the image data. From the CCT-ADDS, the magnitude of the calibration pulse for each scan line could be computed, which in turn allowed calculation of the 'scan-cycle mean' for each scan. The 'rise' for each scan was computed and plotted against the difference between 'scan-cycle mean' and scene mean as illustrated in Figure 5. Regression analysis indicated an excellent fit (R^2 of 0.99), which strongly supports the hypothesis, indicating that the parameter 'B' in the model expressed above is a function of the difference between the 'scan-cycle mean' and the scene mean.

Although this analysis was performed only for forward scans of Band 1 of one Landsat-4 TM image, experience to date indicates that the result may be extended to both scan directions of Bands 1-4 of both Landsat-4 and -5 Thematic Mappers with a high degree of confidence.

This 'droop/rise' effect has been observed for the Primary Focal Plane Bands only. For both Landsats 4 and 5, Bands 5 and 7 show essentially no change in mean signal level within the scan line, with perhaps a slight change in the *opposite* direction to that seen in Bands 1-4. Band 6 mean signal levels have been observed to change within scan lines in a variety of patterns. Detailed analysis of potential within-scan effects in Band 6 is made more difficult by the absence of any constant scene data comparable to the nighttime data in the reflective bands. Even a completely uniform ground scene would have varying atmospheric effects in different parts of the scene.

This 'droop' effect should not cause serious problems for most users. However, it can confound attempts to extend signatures from one side of a scene to the other, and can introduce banding (stripes 16 lines wide, or 17 lines in geometrically corrected CCT-PT data) at the scene edges. Implementation of the proposed exponential model would require pixel-by-pixel correction and could prove costly in terms of computation time. It is our understanding that NASA and NOAA will leave it to the individual users to determine the importance of correcting for this effect and actually performing the correction.

Also apparent in Figures 4a-4h are oscillations superimposed on the 'rise' effect. These oscillations are coherent noise found in all reflective bands of both Landsat-4 and Landsat-5. Although quite obvious in these plots derived from nighttime data, the peak-to-peak amplitudes are quite small (<0.75 DN in unfiltered data, <0.05 DN in these smoothed plots), and have not been observed in daytime data. The cause of this approximately 400 Hz (262-264 pixel period) noise is undetermined.

Scan–Correlated Level Shifts. In the Landsat-4 TM data we have examined, Type 4-7 scan-correlated level shifts are always present, and the signals often shift states with a regular period. Scan-correlated shifts of Type 4-1 are present in most, but not all data, and the Type 4-1 pattern tends to remain in one state or the other for several scans of the scan mirror. The peak-to-peak amplitude for each affected detector for each form of the shift is essentially constant in all cases where that form of the noise exists. The phase relationships between the affected detectors also remain constant in all images (i.e., Band 7 Detector 7 is always in its 'high' Type 4-7 state when Band 5 Detector 8 is in its 'low' Type 4-7 state). Figure 14 of Malila et.al. [1984] illustrated both patterns of level shift for the 16 Band 1 detectors of Landsat-4 TM for a night scene. Relative magnitudes and phases are readily apparent from the illustrations. Table 1 provides the quantitative results, giving the magnitude and phase of each level-shift pattern for the 96 Landsat-4 TM reflective-band detectors.

Initial analyses of Landsat-5 TM data indicated a similar effect, but with only one pattern [Malila and Metzler 1984, Barker 1984]. We examined nighttime reflective-band data to provide quantification of the magnitude and phase

relationships of the effect. Figure 6 illustrates the level shifts for Band 3 of Landsat-5 TM, the band most affected by this noise. The plots were produced by computing the mean signal level for each scan for each detector of each band, and plotting these scan-line means vs the scan number. In these plots, the maximum peak-to-peak amplitude is approximately 0.5 DN. Table 2 contains the quantitative results for the reflective-band detectors of the Landsat-5 TM. It can be seen that nearly all detectors are affected, although the magnitude is very low (<0.1 DN) for many. Band 3 shows the greatest effect, although Band 2 Detector 1 is the single most affected detector with a level shift >0.5 DN. This compares with the maximum shift of 2.0 DN measured for Landsat-4 Band 1 Detector 4. Several detectors did not display any measurable effect in this scene. They are: Band 1 Detectors 1, 3, 5, 9, 13, and 15, Band 2 Detector 4, Band 4 Detectors 8, 10, 12, and 16, Band 5 Detectors 2, 4, 7, 10, and 13, and Band 7 Detectors 1, 2, 5, and 15. As seen in Landsat-4 TM data, patterns of phase and magnitude of the level-shift effect within a band often place the detectors into odd/even groups. As with the within-scan 'droop', the confounding effect of scene data prevents analysis of this type for Band 6. For this band, shutter data may be used to quantify any level shifts, but with somewhat lowered precision.

Although these level-shifts are strikingly evident in the nighttime reflective data, where the scene makes no contribution to the observed signal level, they are of the same magnitude in daytime data and even there can cause noticeable striping. The magnitude of these shifts and the large number of scenes in which they occur places a high value on the correction of the effect for some applications. Fortunately, the constancy of the magnitude permits relatively simple correction techniques. Since the level shift remains constant for the entire scan, the shifts are also observable in the shutter data collected at the ends of each scan line. Based on this, several methods of correcting for level shifts have been proposed which appear effective in reducing the effect [Barker 1984, Fischel 1984, Kogut et.al. 1983, Malila et.al. 1984, Metzler and Malila 1983, Murphy et.al. 1984]. The general approach is to detect the presence of the shift (normally by looking at shutter data), then to subtract (or add) the known magnitude of the shift to each pixel in the affected scan line.

TM Landsat−4 vs Landsat−5 Radiometric Comparison. Radiometric matching of the Landsat-4 and -5 TM sensors was facilitated by the availability of a unique set of radiometrically corrected data collected simultaneously by the two sensors, and registered to sub-pixel accuracy as described above. Same-band images from the two sensors were very similar in appearance, although examination on an image display system required different gain and offset factors to be applied to achieve identical brightness and contrast for each pair of images. Since both the Landsat-4 and Landsat-5 scenes were processed through TIPS (*T*hematic *M*apper *I*mage *P*rocessing *S*ystem), it was expected that radiometrically corrected products would have essentially identical corrected signal values for the same scene viewed at the same time. In addition to multiplicative and additive differences, clipping of the Landsat-5 data values was obvious in both Bands 5 and 7 at the low radiance end of the dynamic range. The Band 7 low-level clipping is apparent from a histogram of signal-level frequency for Band 7 for both Landsat-4 and Landsat-5 (see Figure 7). The pixels with values zero to six in the Landsat-4 scene are all mapped to value zero in the Landsat-5 scene. Although the offset was nearly as large for

Band 5 (see Figure 8), fewer data values actually were clipped (0.3 percent of the scene versus 4.2 percent in Band 7).

As noted earlier, band-by-band comparisons were carried out using two different techniques; regression of signal values from the coincident pixels or regions, and regression of signal values associated with specific histogram percentile classes. When clipping was not present, either technique produced essentially the same results. Where clipping was present, regression of matched areas led to smaller additive terms and larger multiplicative terms, a result deemed erroneous after inspecting the histograms. For this reason, coefficients from the histogram matching approach are presented here. Table 3 presents the multiplicative and additive coefficients to convert Landsat-4 TM signal levels to Landsat-5 TM equivalent values; Table 4 contains the coefficients to convert Landsat-5 signals to Landsat-4 values. It should be noted that while this was a simultaneously collected data set, and therefore nearly ideal for this type of analysis, the correction coefficients presented are valid only if ground processing parameters are not changed. It should also be noted that data which have been clipped as in Landsat-5 Bands 5 and 7 can not be retrieved — all the zeroes in Landsat-5 Band 7 data will be converted to sixes in Landsat-4 Band 7 whereas Landsat-4 Band 7 would have recorded the same pixels with signal levels ranging from zero to six. In using these conversion equations, resultant DN's less than zero should be assigned the value zero; DN's greater than 255 should be assigned 255.

Converting the pixel values to radiance levels via the coefficients provided in the Radiometric Calibration Ancillary Record of the Leader File associated with each band of image data [NASA 1983] did not resolve the discrepancy observed between the two sensors. Table 5 lists the multiplicative and additive coefficients extracted from tape headers and used in the conversion; Table 6 is similar to Table 3 in that it defines conversion of Landsat-4 signals to Landsat-5 equivalent signals, but in terms of radiance instead of signal counts. It is not known at this time why the radiometrically corrected data are not more closely matched.

An additional discrepancy was noted between the previously published Band 6 temperature sensitivity range and the range implied by the coefficients listed in Table 5. Using these coefficients to convert the range 0-255 DN to radiance gives a radiance range of 0.125 to 1.575 mW/cm^2-sr-um, representing an apparent temperature range of approximately 200 to 340°Kelvin, not the advertised 260°K to 320°K. This causes an increase in the temperature difference represented by a change of 1 DN. The specified 260°K to 320°K temperature range actually spans approximately 63-196 DN vs the specified 0-255 DN. For Landsat-5 TM, the radiance range is very slightly different (0.124 to 1.560 mW/cm^2-sr-um), still giving a range of apparent temperature of approximately 200°K to 340°K (or a DN range of approximately 63-193 for apparent temperatures of 260°K to 320°K). Users unaware of these differences may incorrectly derive temperatures from TM Band 6 data.

4. SUMMARY

Landsat-5 TM image data were found to be quite similar to Landsat-4 TM data, both in terms of high overall quality and in the presence of several anomalies. Detailed analysis revealed a systematic within-scan drift (or 'droop/rise') of the signal from the scene mean toward the overall 'scan-cycle mean' in spectral Bands 1-4. The magnitude of this drift ranged from minus 1.5 DN (daytime) to +0.15 DN (nighttime), depending on scene content. The drift was fitted with a simple exponential decay model and found to have a time constant equivalent to about one-sixth of a frame width.

Scan-correlated level shifts are present in both Landsat-4 and Landsat-5 TM data. The maximum effect observed in Landsat-5 data was approximately 0.5 DN peak-to-peak, compared with a maximum of 2.0 DN observed in Landsat-4 data. The level-shifts appear to be present in most if not all images, and effective correction procedures have been proposed.

Although data from both Thematic Mappers are produced in radiometrically corrected form, comparison of data acquired simultaneously by the two sensors revealed significant differences in their calibration. In the reflective bands, the multiplicative factors required to convert Landsat-4 TM data to Landsat-5 data ranged from 0.987 to 1.145, with corresponding additive terms of −2.7 DN to −6.2 DN, and displayed evidence of low-level clipping in Landsat-5 Bands 5 and 7. The thermal bands (Band 6) were more closely matched, but are calibrated to have a full-range temperature range of 200°K to 340°K instead of the advertised 260°K to 320°K.

REFERENCES

Barker, J.L., "Relative Radiometric Calibration of Landsat TM Reflective Bands," *LANDSAT−4 Science Investigations Summary*, vol. I, pp. 140-180, July 1984.

Barker, J.L., R.B. Abrams, D.L. Ball, and K.C. Leung, "Radiometric Calibration and Processing Procedure for Reflective Bands on Landsat-4 Protoflight Thematic Mapper", *Proc. of Landsat−4 Science Characterization Early Results Symp.*, vol. II, p.75, January 1985.

Fischel, D., "Validation of the Thematic Mapper Radiometric and Geometric Correction Algorithms," *IEEE Trans. Geosc. Remote Sensing*, vol. GE-22, no.3, pp. 237-242, May 1984.

Kieffer, H., D.A. Cook, E.M. Eliason, P.T. Eliason, "Intraband Radiometric Performance of the Landsat Thematic Mappers," *Photogrammetric Eng. and Remote Sensing*, this issue.

Kogut, J., E. Larduinat, and M. Fitzgerald, "An Analysis of New Techniques for Radiometric Correction of LANDSAT-4 Thematic Mapper Images," Research and Data Systems, Inc., Lanham, MD, October 1983.

Malila, W.A., M.D. Metzler, D.P. Rice, and E.P. Crist, "Characterization of LANDSAT-4 MSS and TM Digital Image Data," *IEEE Trans. Geosc. Remote Sensing*, vol. GE-22, no.3, pp. 177-191, May 1984.

Metzler, M.D., and W.A. Malila, "Radiometric Characterization of Thematic Mapper Full-frame Imagery," in *Proc. SPSE/ASP Conf. on Techniques for Extraction of Information from Remotely Sensed Images*, Rochester, NY, 1983.

Metzler, M.D., and W.A. Malila, "Scan-angle and Detector Effects in Thematic Mapper Radiometry," *Proc. of Landsat−4 Science Characterization Early Results Symp.*, vol. II, pp.421-441, January 1985.

Malila, W.A., and M.D. Metzler, *Eighth Quarterly Progress Report for Study of Spectral/Radiometric Characteristics of the Thematic Mapper for Land Use Applications*, ERIM Report No. 164000-13-P, October 1984.

Murphy, J.M., T. Butlin, P.F. Duff, and A.J. Fitzgerald, "Revised Radiometric Calibration Technique for LANDSAT-4 Thematic Mapper Data," *IEEE Trans. Geosc. Remote Sensing*, vol. GE-22, no.3, pp. 243-251, May 1984.

NASA, "Interface Control Document between the NASA Goddard Space Flight Center (GSFC) and the Department of Interior EROS Data Center (EDC) for Landsat-D Thematic Mapper Computer Compatible Tape (CCT-AT, CCT-PT), Revision A," LSD-ICD-105, pp. 60, December 1983.

TABLE 1. Magnitude of Level Shifts for Night Scene 4-0161-02481 (Landsat-4)

Band	Det	Amplitude Form 1	Amplitude Form 2	Separation of States (#S.D.) Form 1	Separation of States (#S.D.) Form 2
1	1	0.261	-0.162	3.1	1.9
1	2	-0.024	-0.188	.3	2.5
1	3	0.077	-0.117	1.1	1.7
1	4	1.880	0.116	26.2	1.6
1	5	0.120	-0.141	2.1	2.4
1	6	0.038	-0.060	.5	.8
1	7	0.098	-0.077	1.6	1.2
1	8	0.569	-0.092	8.0	1.3
1	9	0.146	0.043	2.0	.6
1	10	0.877	-0.053	13.2	.8
1	11	0.173	-0.063	2.1	.8
1	12	1.601	-0.011	24.3	.2
1	13	0.313	-0.029	3.8	.3
1	14	0.084	-0.124	1.2	1.8
1	15	0.171	-0.041	1.8	.4
1	16	-0.025	-0.123	.3	1.6
2	1	0.248	-0.323	4.0	5.2
2	2	0.149	0.056	3.0	1.1
2	3	-0.082	-0.202	1.2	3.0
2	4	0.203	0.027	5.0	.7
2	5	0.072	-0.138	1.4	2.7
2	6	0.075	0.056	1.4	1.1
2	7	0.004	-0.086	.1	2.0
2	8	0.106	0.062	1.8	1.1
2	9	0.039	-0.053	1.3	1.8
2	10	0.033	0.031	1.3	1.2
2	11	0.056	-0.061	1.4	1.5
2	12	0.143	0.035	3.7	.9
2	13	0.084	-0.101	1.7	2.1
2	14	0.064	0.043	1.7	1.2
2	15	0.097	-0.099	2.5	2.6
2	16	0.074	0.127	1.8	3.1
3	1	0.484	-0.494	7.3	7.5
3	2	0.296	0.273	7.3	6.7
3	3	0.235	-0.275	4.9	5.3
3	4	0.022	0.032	.7	.9
3	5	0.304	-0.205	6.0	4.0
3	6	0.248	0.086	5.7	2.0
3	7	0.155	-0.203	3.6	4.8
3	8	-0.088	-0.034	2.1	.8
3	9	-0.051	-0.063	1.2	1.4
3	10	0.253	0.055	4.0	.9
3	11	0.003	-0.134	.2	4.0
3	12	0.102	0.230	1.9	4.2
3	13	0.215	-0.151	5.6	3.9
3	14	0.209	0.180	4.2	3.6
3	15	0.278	-0.262	6.9	6.5
3	16	0.446	0.782	6.9	12.1
4	1	0.256	-0.047	6.6	1.2
4	2	0.147	-0.084	3.3	1.9
4	3	0.197	-0.031	2.3	.4
4	4	0.052	-0.072	1.3	1.8
4	5	0.119	-0.013	3.7	.4
4	6	0.316	-0.063	5.2	1.0
4	7	0.073	-0.017	3.1	.7
4	8	0.003	-0.083	.1	2.4
4	9	0.064	-0.023	1.9	.7
4	10	0.147	-0.013	5.0	.4
4	11	0.101	-0.017	4.0	.6
4	12	0.252	0.019	7.0	.5
4	13	0.063	-0.008	4.5	.6
4	14	0.087	-0.020	3.2	.7
4	15	0.114	-0.015	5.2	.6
4	16	0.371	0.029	9.7	.7
5	1	-0.018	0.060	.6	2.2
5	2	0.136	-0.105	3.6	2.8
5	3	-0.125	-0.067	4.6	2.5
5	4	0.113	-0.038	3.3	1.1
5	5	-0.183	-0.043	6.6	1.5
5	6	0.046	-0.085	1.4	2.6
5	7	0.017	0.143	.5	4.8
5	8	-0.074	-0.200	2.6	6.9
5	9	-0.085	0.023	2.7	.7
5	10	-0.086	0.744	2.7	22.7
5	11	-0.022	0.046	.9	1.9
5	12	0.006	0.061	.3	2.6
5	13	-0.083	0.004	3.4	.2
5	14	-0.079	-0.082	3.8	3.9
5	15	-0.211	-0.119	8.3	4.6
5	16	-0.017	-0.100	.7	4.3
7	1	-0.034	0.302	1.3	11.5
7	2	0.054	-0.293	1.8	9.7
7	3	-0.090	0.342	3.1	11.8
7	4	0.093	-0.273	3.4	9.9
7	5	-0.006	0.281	.2	10.8
7	6	0.123	-0.308	4.3	10.4
7	7	-0.040	0.960	.3	6.4
7	8	0.066	-0.345	2.6	13.5
7	9	-0.109	0.197	4.2	7.5
7	10	0.095	-0.461	3.3	15.1
7	11	-0.045	0.232	1.6	8.5
7	12	0.168	-0.297	5.3	9.1
7	13	-0.087	0.224	2.8	7.1
7	14	0.198	-0.085	6.6	2.8
7	15	-0.043	0.193	1.7	7.7
7	16	0.169	-0.253	5.5	8.2

** Negative amplitudes indicate level shifts with phase shifts of 180° relative to B1D4 (Form 1) or B7D7 (Form 2).

ERIM 6 Dec 83

TABLE 2. Magnitude of Level Shifts for Night Scene 5-0052-02182 (Landsat 5)

Band	Det	Amplitude	Separation of States (Number of Std. Dev.)	Band	Det	Amplitude	Separation of States (Number of Std. Dev.)
1	1	0.036	0.9	4	1	0.285	11.0
1	2	-0.229	6.4	4	2	0.150	5.5
1	3	0.061	1.7	4	3	0.149	5.0
1	4	-0.184	4.2	4	4	0.205	8.4
1	5	0.044	1.3	4	5	0.044	3.3
1	6	-0.249	6.3	4	6	0.045	2.5
1	7	0.113	3.4	4	7	0.025	3.1
1	8	-0.208	5.2	4	8	0.016	0.6
1	9	0.042	1.3	4	9	0.078	2.8
1	10	-0.303	9.3	4	10	0.004	0.3
1	11	0.068	2.3	4	11	0.106	3.8
1	12	-0.216	5.4	4	12	0.016	0.7
1	13	-0.010	0.3	4	13	0.138	4.4
1	14	-0.307	8.5	4	14	0.081	2.2
1	15	0.001	0.0	4	15	0.085	3.0
1	16	-0.266	7.6	4	16	0.015	0.5
2	1	0.523	16.3	5	1	0.135	6.8
2	2	0.084	4.6	5	2	0.006	0.4
2	3	0.239	11.5	5	3	0.190	13.6
2	4	-0.009	1.1	5	4	-0.023	1.7
2	5	0.178	10.5	5	5	-0.067	4.5
2	6	0.070	7.2	5	6	-0.128	9.3
2	7	0.157	8.7	5	7	0.008	0.6
2	8	0.065	7.3	5	8	-0.115	9.0
2	9	0.079	8.2	5	9	0.056	4.3
2	10	0.012	2.4	5	10	0.019	0.5
2	11	0.086	6.9	5	11	0.100	7.4
2	12	0.016	2.3	5	12	-0.129	6.9
2	13	0.177	9.5	5	13	0.022	1.7
2	14	0.075	8.4	5	14	-0.074	5.0
2	15	0.159	10.1	5	15	-0.088	7.4
2	16	0.263	12.6	5	16	-0.137	9.7
3	1	0.470	14.1	7	1	0.019	1.1
3	2	0.196	8.1	7	2	-0.013	0.8
3	3	0.437	14.3	7	3	-0.050	3.1
3	4	0.389	13.9	7	4	-0.110	6.8
3	5	0.317	8.2	7	5	0.001	0.1
3	6	0.340	12.0	7	6	-0.095	6.6
3	7	0.286	7.8	7	7	0.146	8.9
3	8	0.141	8.5	7	8	-0.159	10.2
3	9	0.326	10.5	7	9	0.161	11.7
3	10	0.250	9.4	7	10	-0.156	9.4
3	11	0.371	12.9	7	11	0.107	7.0
3	12	0.351	11.8	7	12	-0.060	4.1
3	13	0.318	12.8	7	13	0.032	2.3
3	14	0.237	7.7	7	14	-0.038	2.1
3	15	0.321	12.6	7	15	0.022	1.6
3	16	0.257	11.8	7	16	-0.116	8.0

** Negative amplitudes indicate level shifts with phase shifts of 180° relative to Band 3 detectors.

TABLE 3. Coefficients for Converting Landsat-4 Values to Landsat-5 Values (Scenes 4-0608-15463 and 5-0014-15460, 15 March 1984)

Landsat-5 TM = A*(Landsat-4 TM) + B

Band	A	B	S.E.	R^2	Range of Data Values Landsat-4	Landsat-5
1	1.0438	-3.538	0.151	0.99943	73-109	73-111
2	1.1200	-2.719	0.134	0.99922	26-52	26-56
3	0.9869	-3.678	0.142	0.99975	26-77	22-72
4	1.0030	-4.627	0.078	0.99995	11-92	7-88
5	1.1452	-7.330	0.106	0.99999	6-154	0-169
6	1.0040	-0.711	0.119	0.99956	114-148	113-148
7	1.0923	-6.244	0.054	0.99999	3-86	0-88

Note: If the value computed for Landsat-5 is <0, substitute 0.
If it is >255, substitute 255.

TABLE 4. Coefficients for Converting Landsat-5 Values to Landsat-4 Values (Scenes 4-0608-15463 and 5-0014-15460, 15 March 1984)

Landsat-4 TM = A*(Landsat-5 TM) + B

Band	A	B	S.E.	R^2	Range of Data Values Landsat-4	Landsat-5
1	0.9580	3.390	0.145	0.99943	73-109	73-111
2	0.8928	2.427	0.120	0.99922	26-52	26-56
3	1.0132	3.726	0.144	0.99975	26-77	22-72
4	0.9970	4.614	0.078	0.99995	11-92	7-88
5	0.8732	6.401	0.093	0.99999	6-154	0-169
6	0.9960	0.714	0.118	0.99956	114-148	113-148
7	0.9155	5.717	0.049	0.99999	3-86	0-88

Note: If the value computed for Landsat-4 is >255, substitute 255.

TABLE 5. Landsats-4 and 5 TM Radiance Conversion Parameters (Scenes 4-0608-15463 and 5-0014-15460, 15 March 1984)

Radiance = $A0 + A1 * DN$ (mW/cm²-sr-um)

Band	$A0$ (mW/cm²-sr-um) Landsat-4	$A0$ (mW/cm²-sr-um) Landsat-5	$A1$ (mW/cm²-sr-um)/DN Landsat-4	$A1$ (mW/cm²-sr-um)/DN Landsat-5
1	-0.1500	-0.1500	0.06024	0.06024
2	-0.2802	-0.2805	0.11750	0.11750
3	-0.1203	-0.1194	0.08061	0.08059
4	-0.1504	-0.1500	0.08145	0.08143
5	-0.0372	-0.0370	0.01081	0.01081
6	0.1252	0.1238	0.00569	0.00563
7	-0.1500	-0.1500	0.00570	0.00568

TABLE 6. Landsats-4 and 5 TM Regressions of Radiance Values (Scenes 4-0608-15463 and 5-0014-15460, 15 March 1984)

Landsat-5 TM = A*(Landsat-4 TM) + B

Band	A	B	S.E.	R^2	Range of Radiance Values (mW/cm²-sr-um) Landsat-4	Range of Radiance Values (mW/cm²-sr-um) Landsat-5
1	1.0435	-0.205	0.009	0.99943	4.25-6.44	4.23-6.52
2	1.1196	-0.285	0.016	0.99922	2.76-5.87	2.79-6.28
3	0.9865	-0.297	0.011	0.99975	1.98-6.06	1.67-5.68
4	1.0027	-0.376	0.006	0.99995	0.77-7.33	0.42-6.98
5	1.1452	-0.074	0.001	0.99999	-.03-1.63	-.03-1.79
6	0.9932	-0.002	0.001	0.99956	0.77-0.97	0.76-0.96
7	1.0885	-0.002	0.001	0.99999	-.13-0.34	-.15-0.35

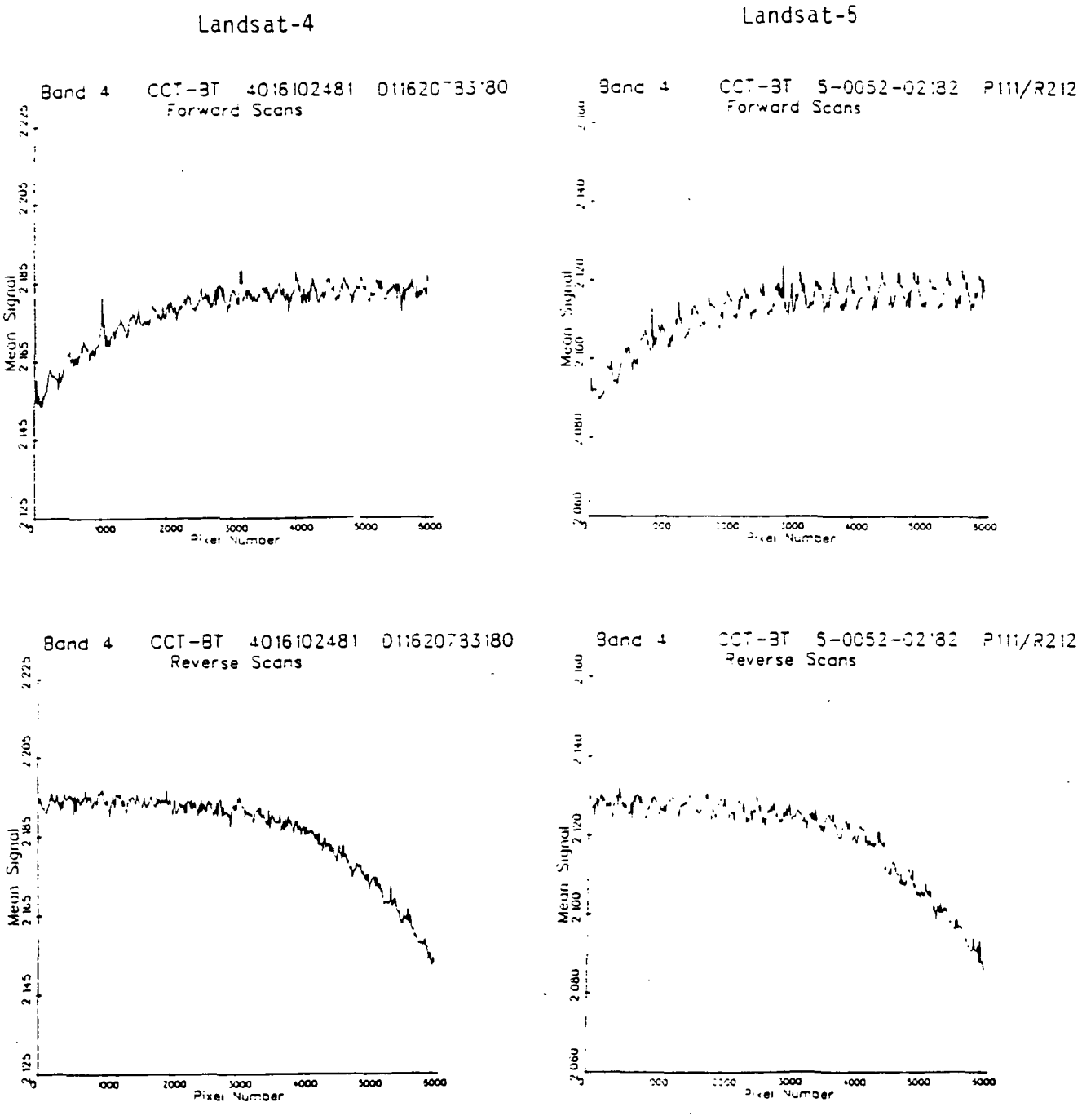


Figure 1. Landsat-4 and -5 TM Nighttime Within-Scan Droop Effect - Band 1

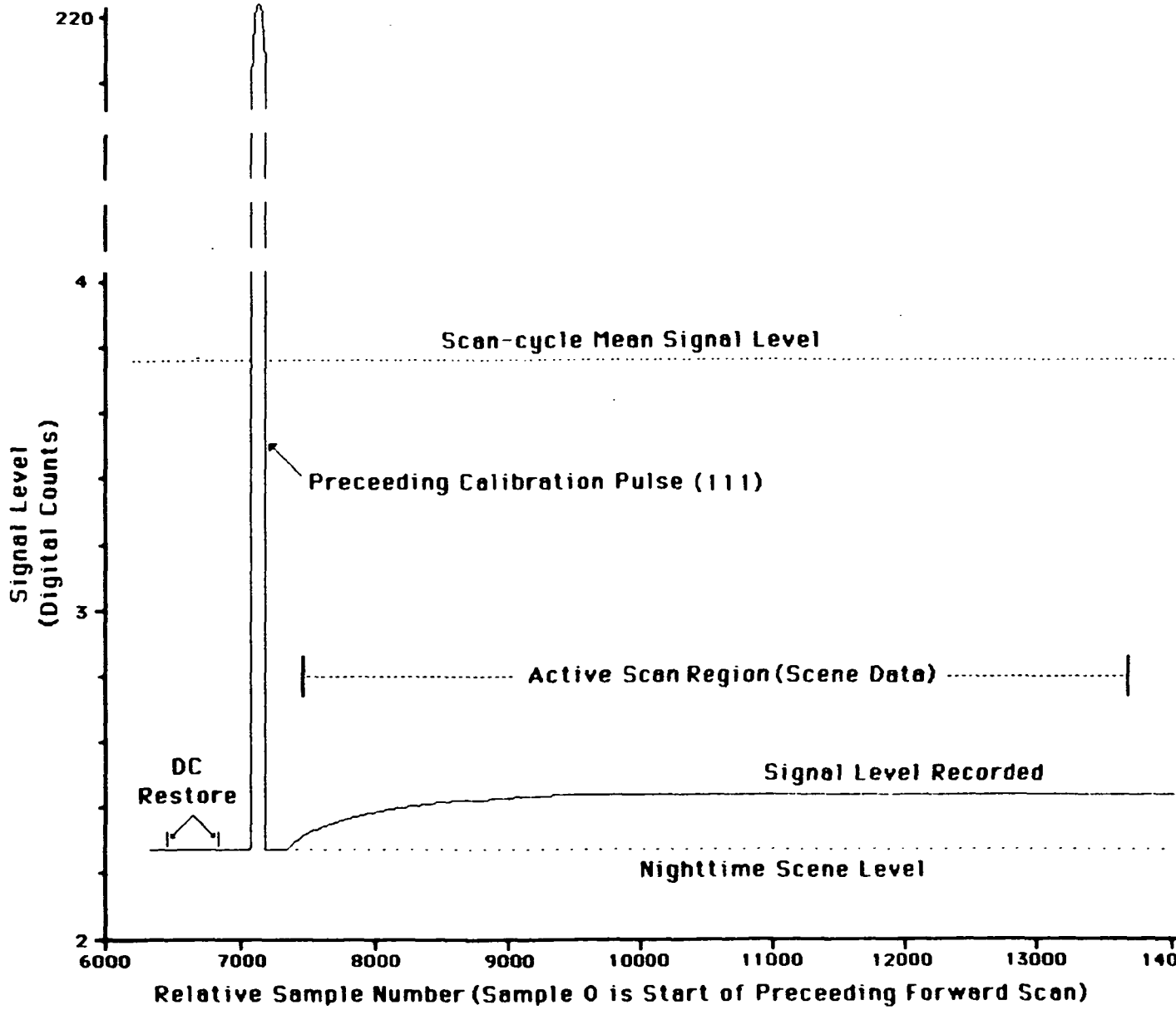


Figure 2. Example Nighttime Reverse Scan Signal Levels — Band 1

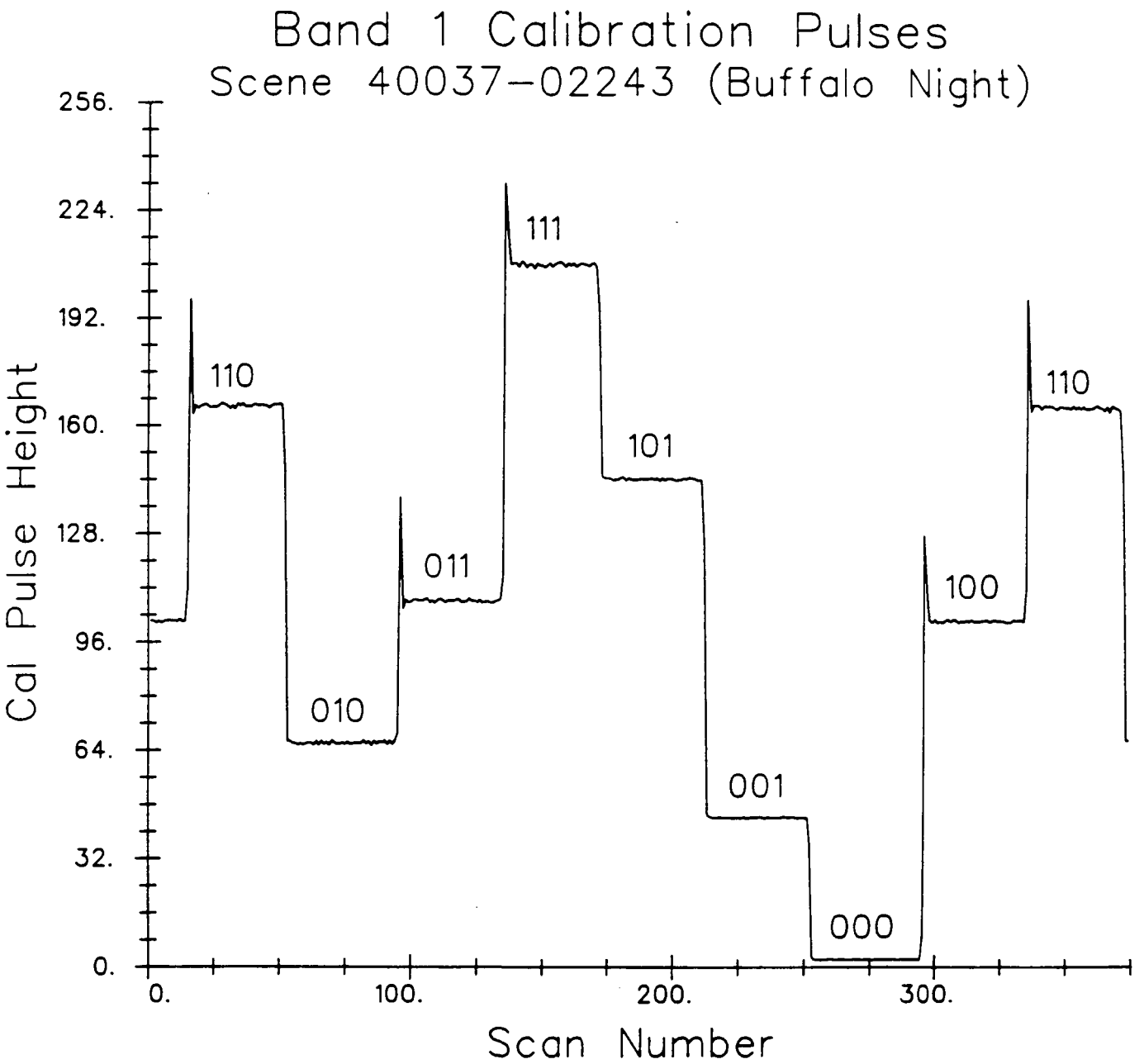


Figure 3. Calibration Lamp Sequencing for Band 1

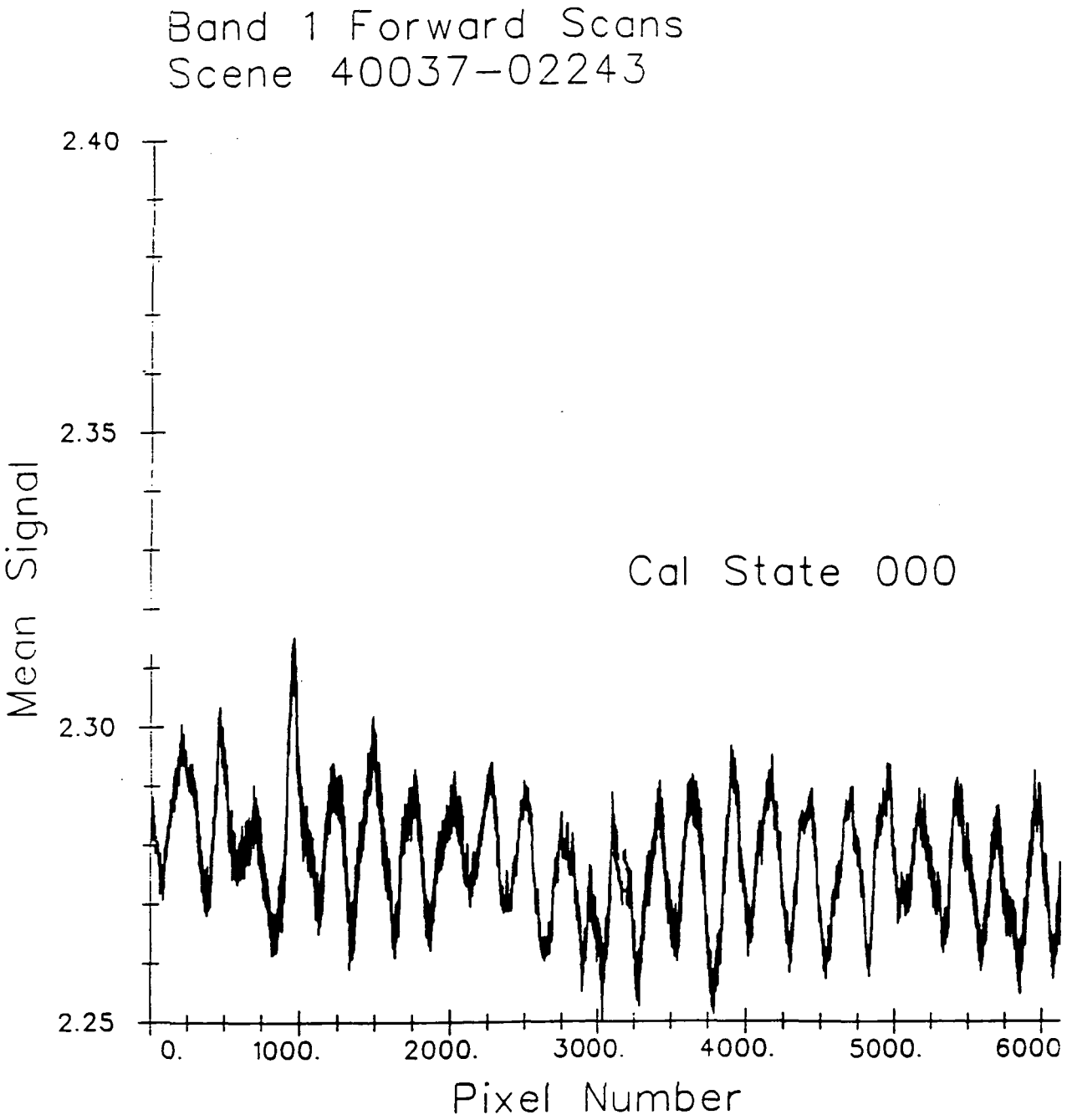


Figure 4a. Nighttime Forward Scan Signal Rise for Scans Preceeded by Calibration Lamp State 000

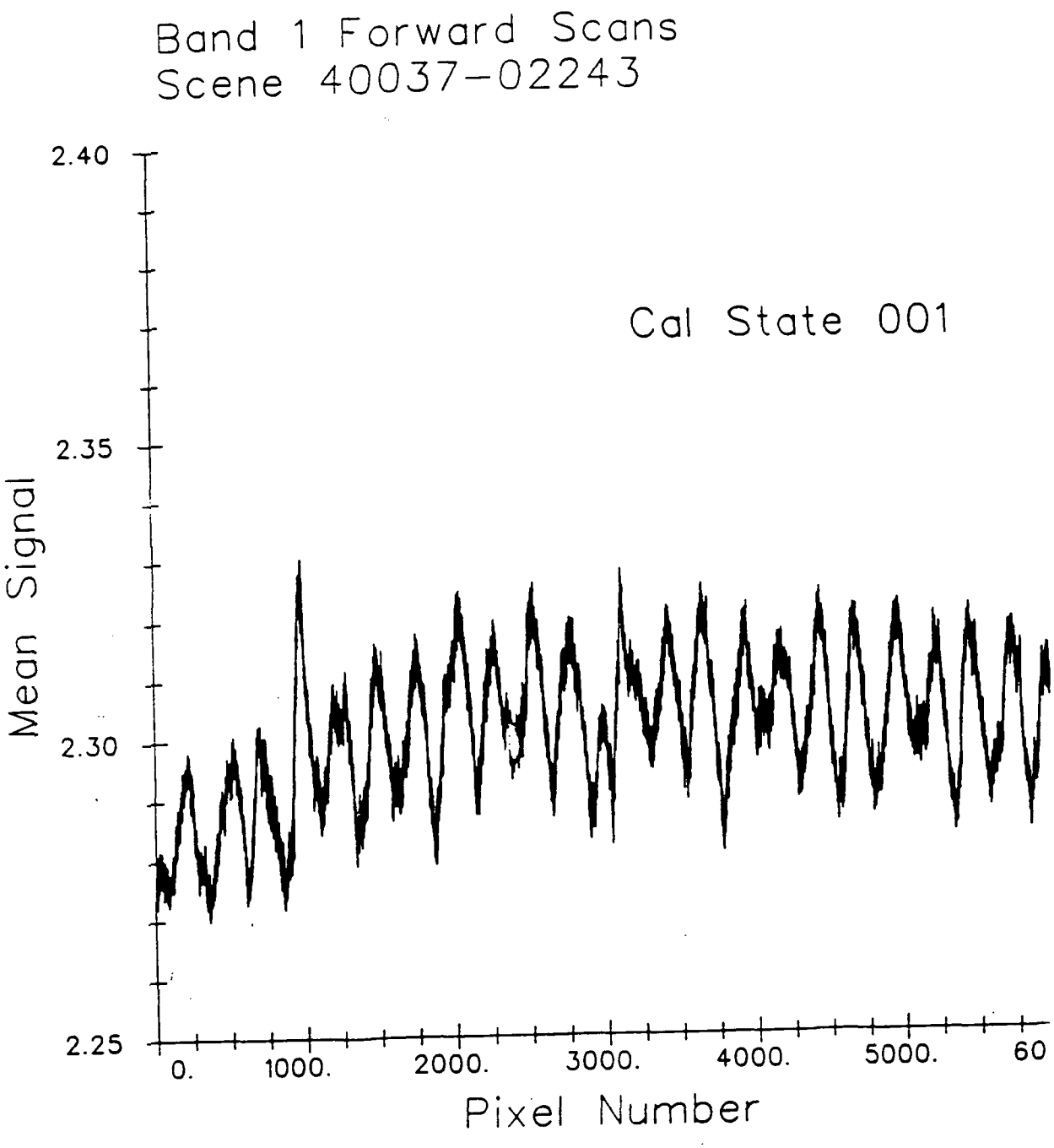


Figure 4b. Nighttime Forward Scan Signal Rise for Scans Preceeded by Calibration Lamp State 001

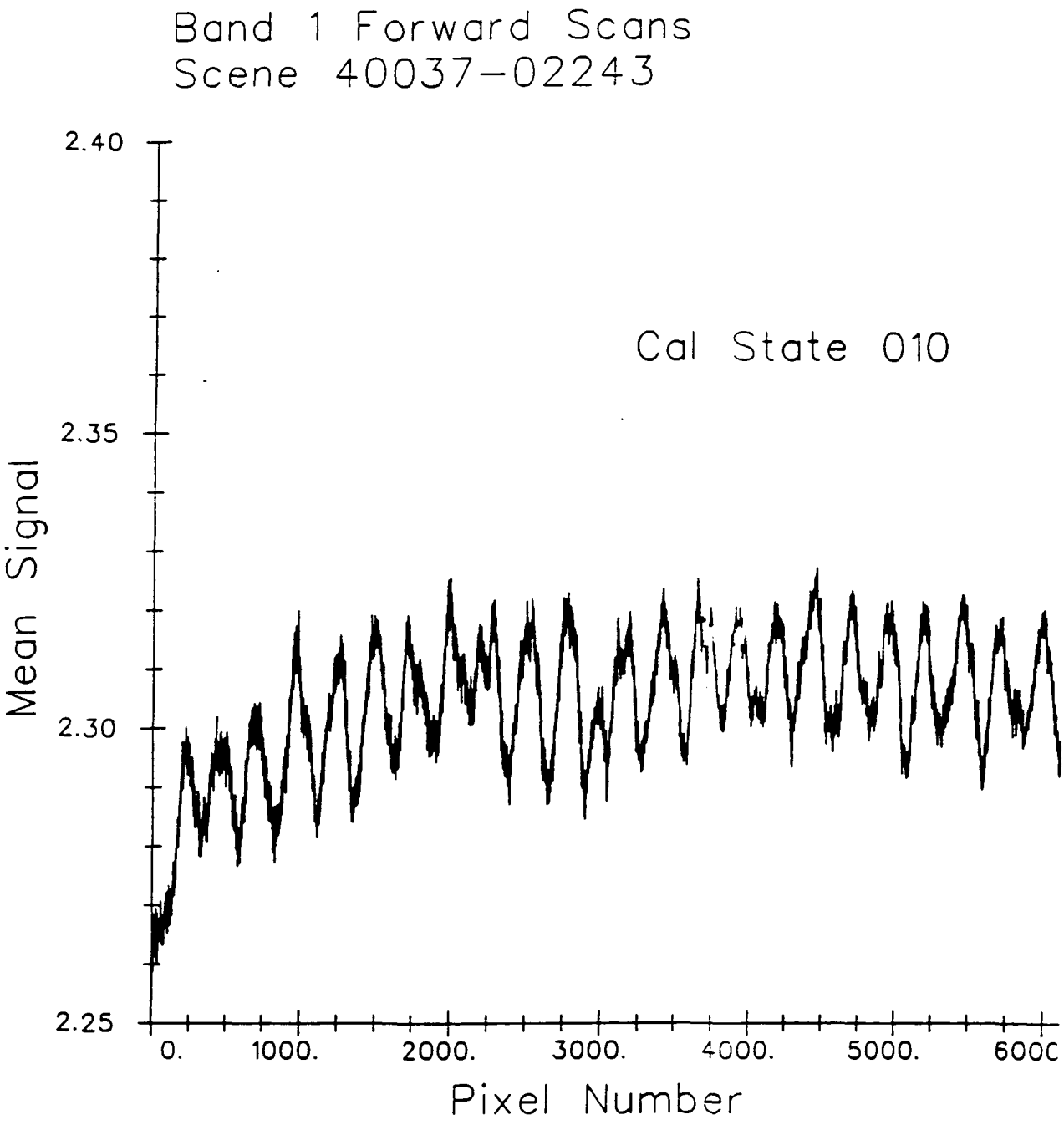


Figure 4c. Nighttime Forward Scan Signal Rise for Scans Preceeded by Calibration Lamp State 010

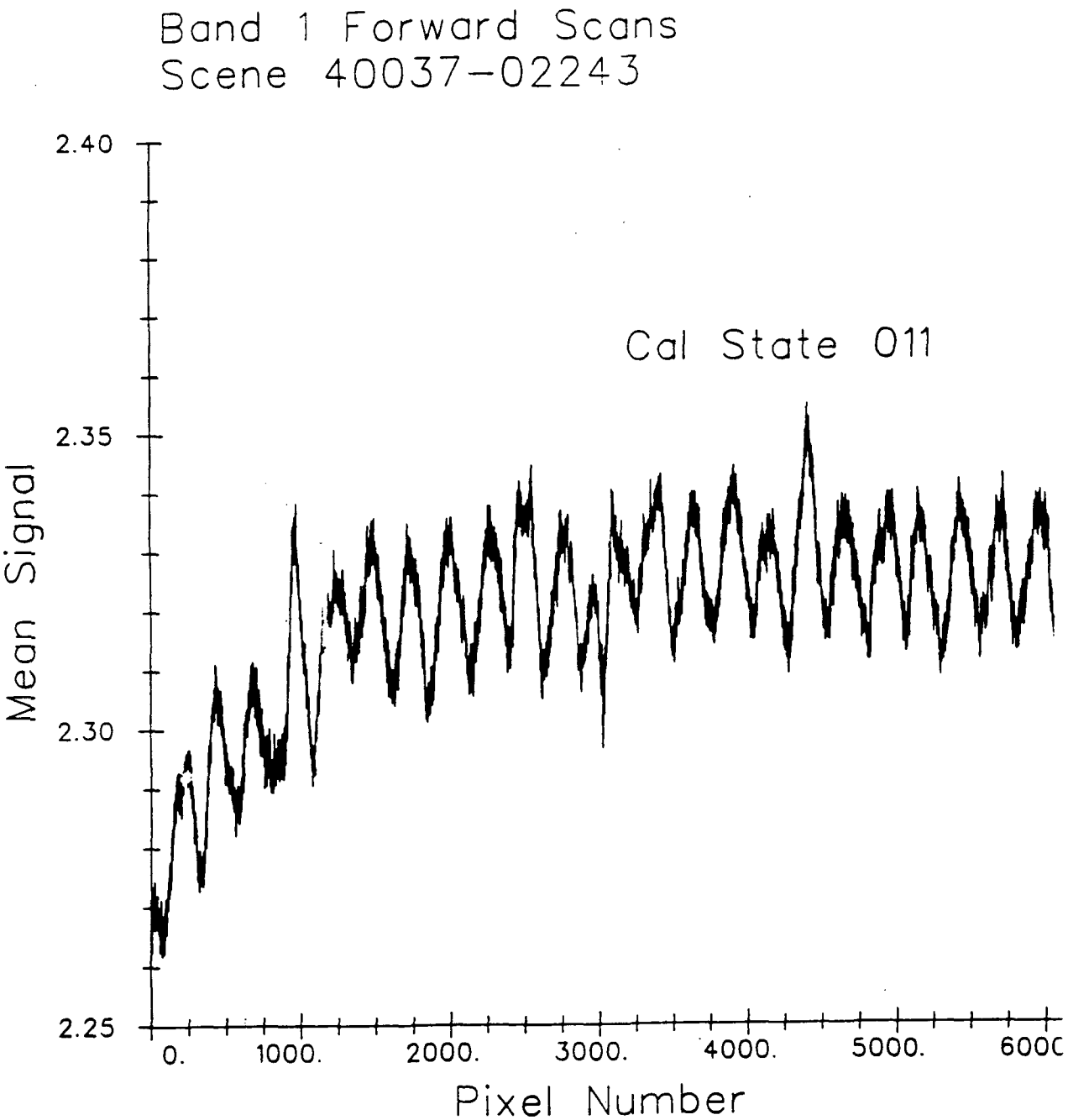


Figure 4d. Nighttime Forward Scan Signal Rise for Scans Preceeded by Calibration Lamp State 011

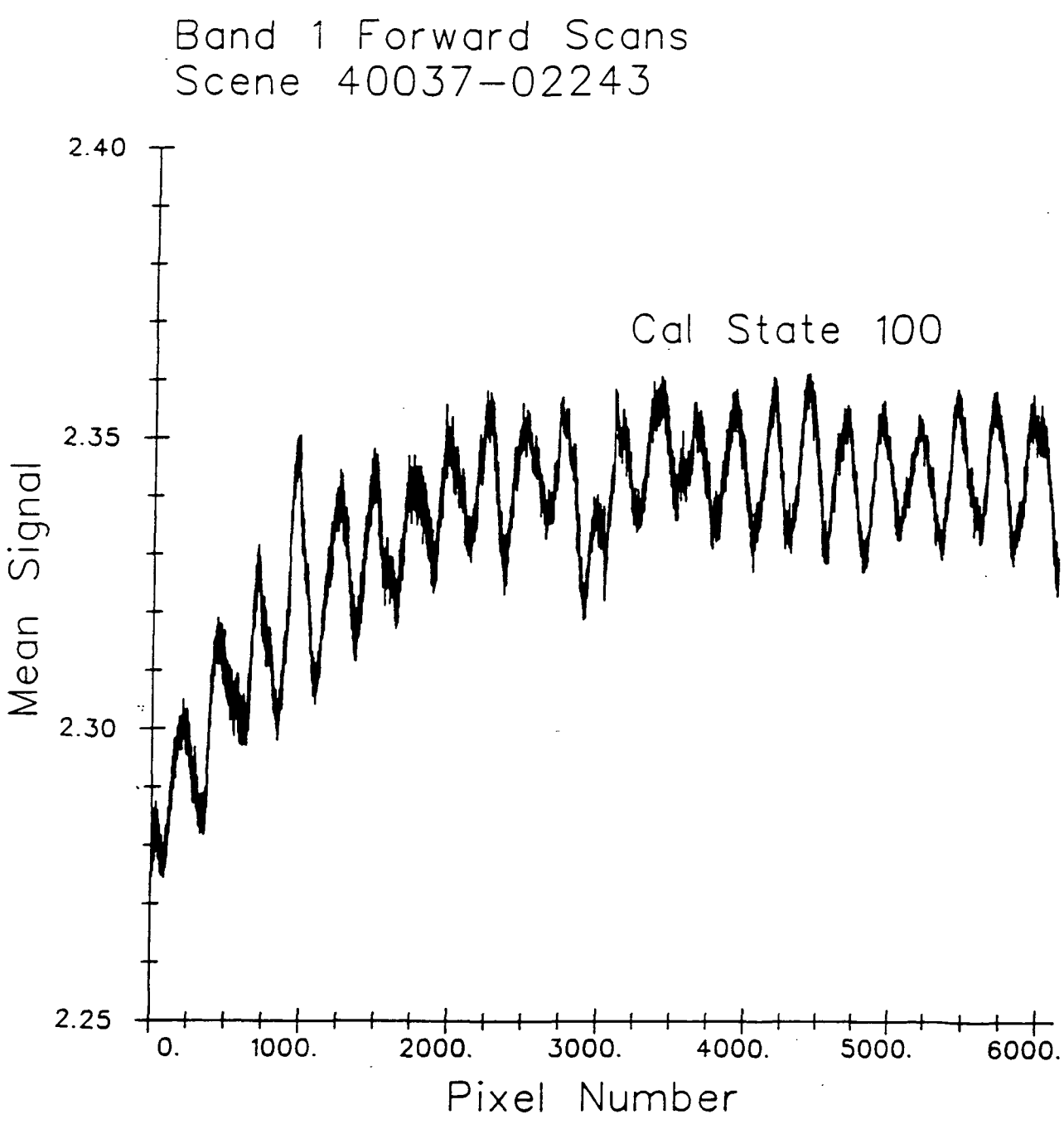


Figure 4e. Nighttime Forward Scan Signal Rise for Scans Preceeded by Calibration Lamp State 100

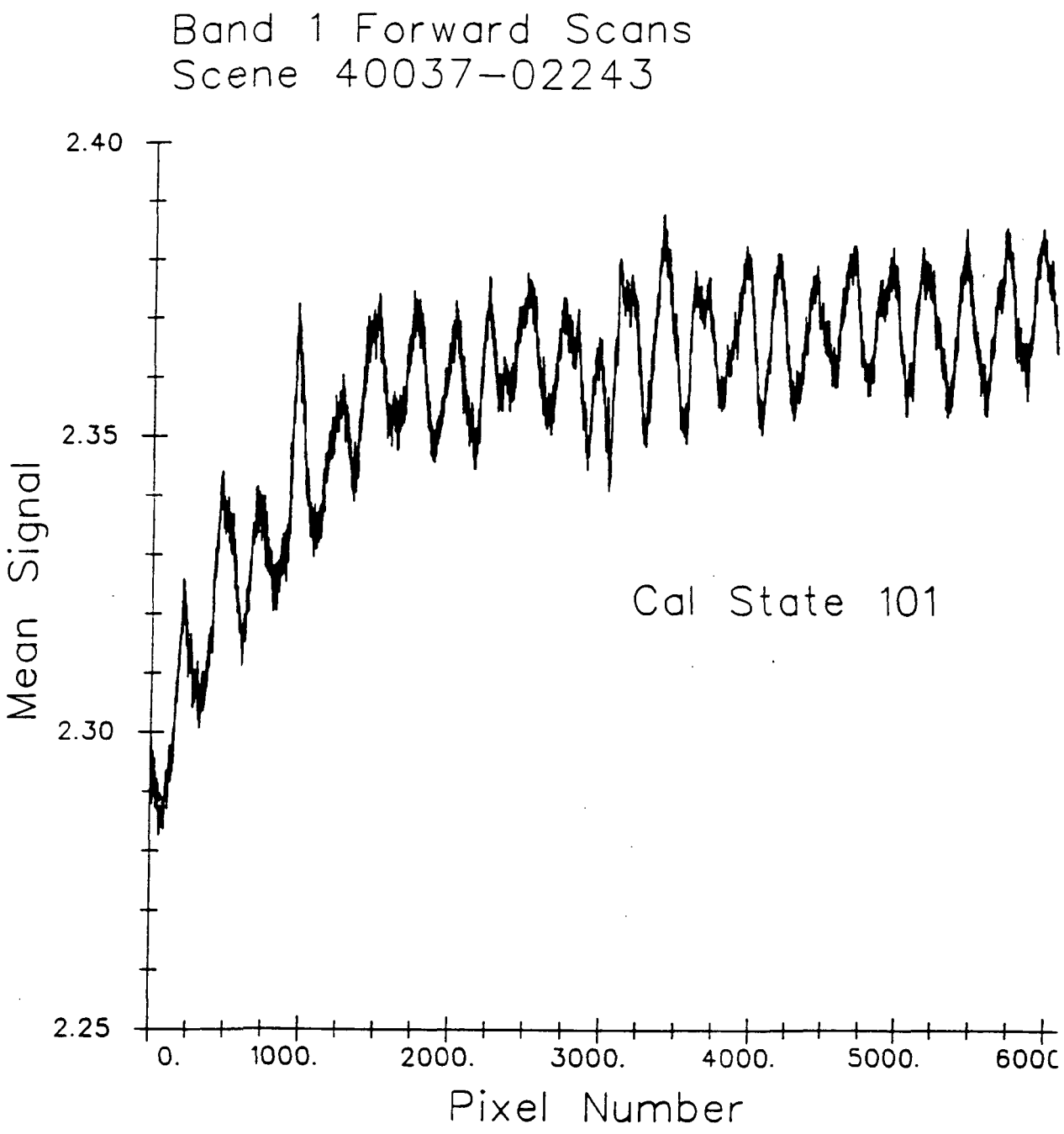


Figure 4f. Nighttime Forward Scan Signal Rise for Scans Preceeded by Calibration Lamp State 101

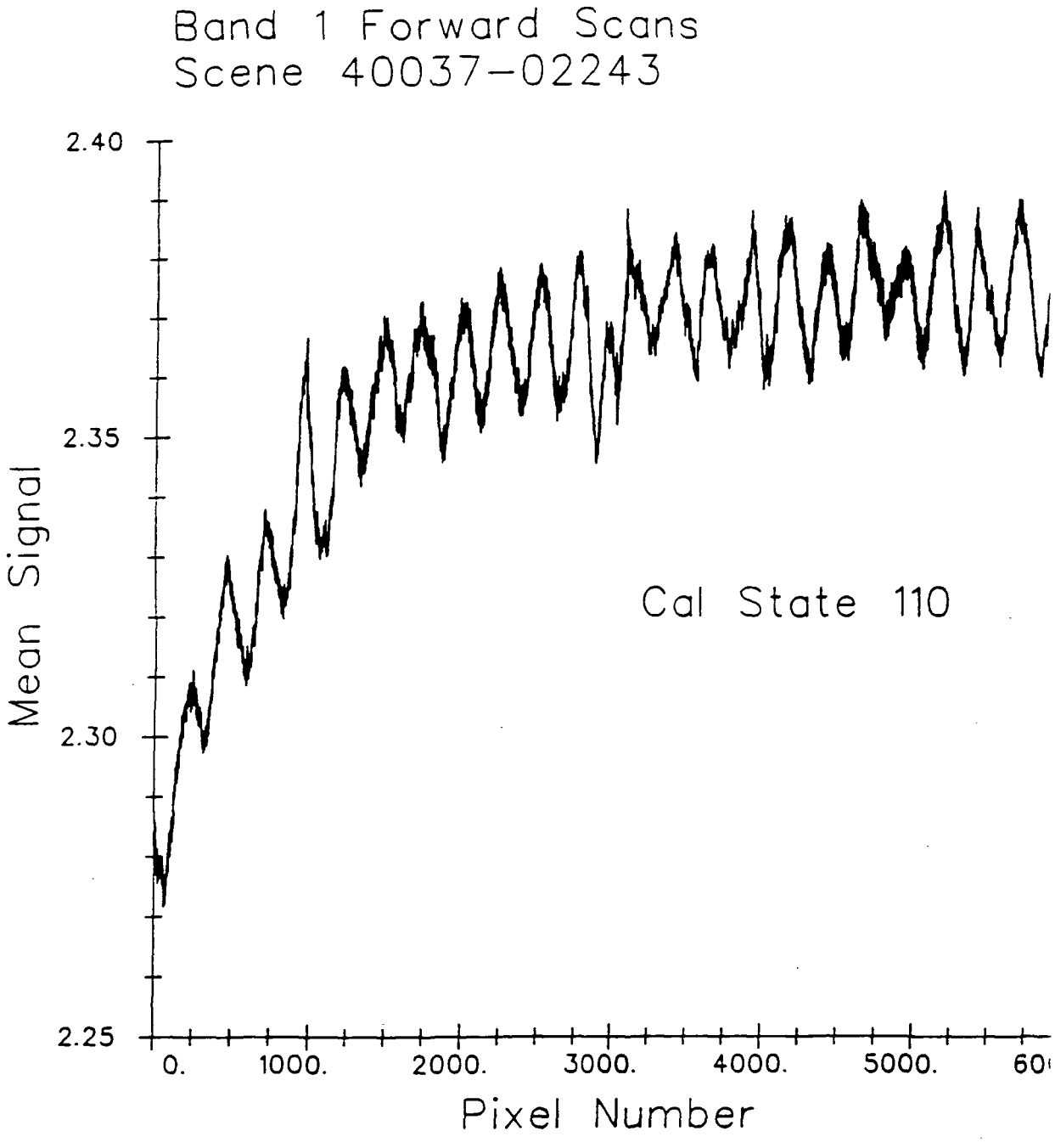


Figure 4g. Nighttime Forward Scan Signal Rise for Scans Preceeded by Calibration Lamp State 110

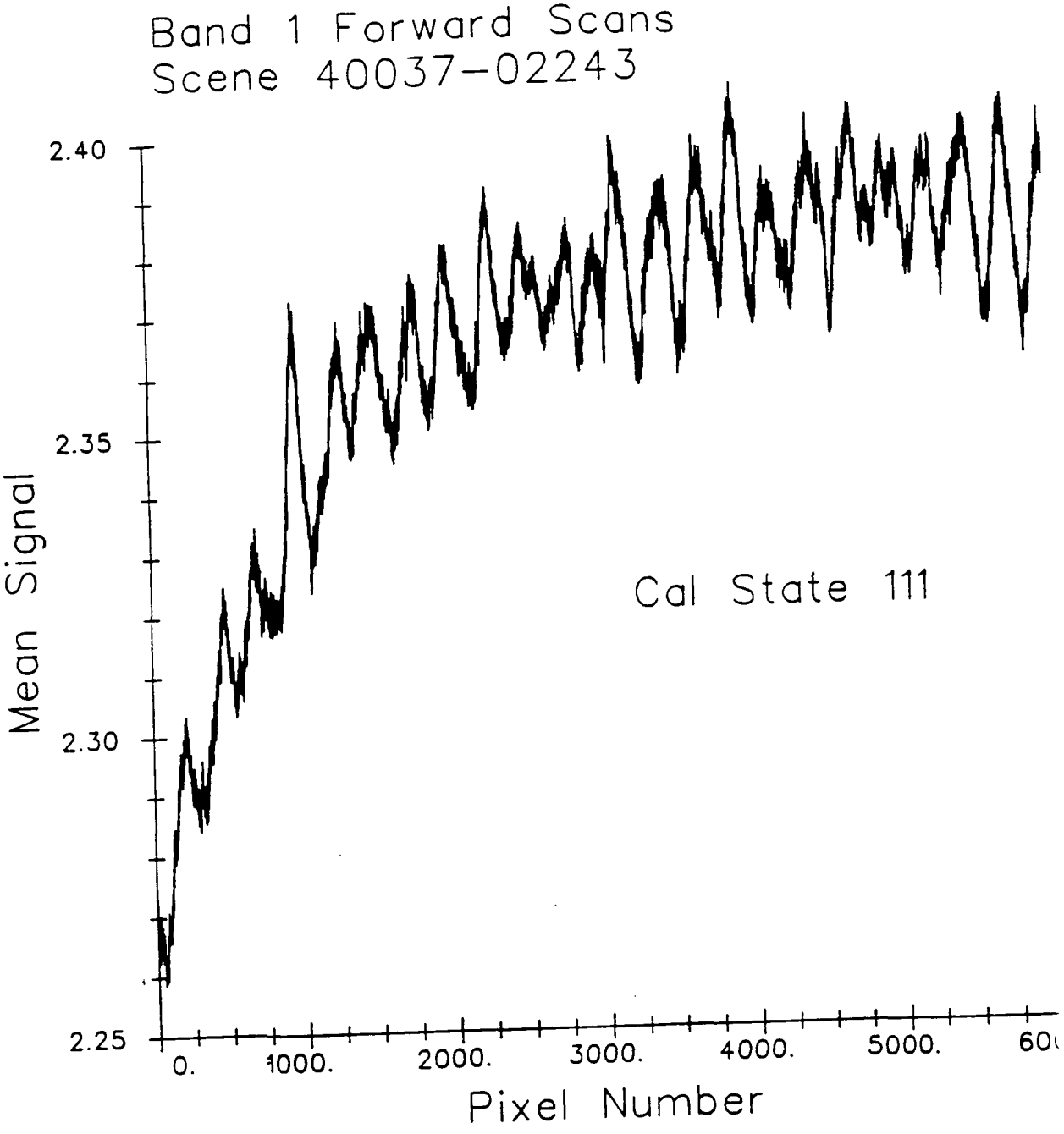


Figure 4h. Nighttime Forward Scan Signal Rise for Scans Preceeded by Calibration Lamp State 111

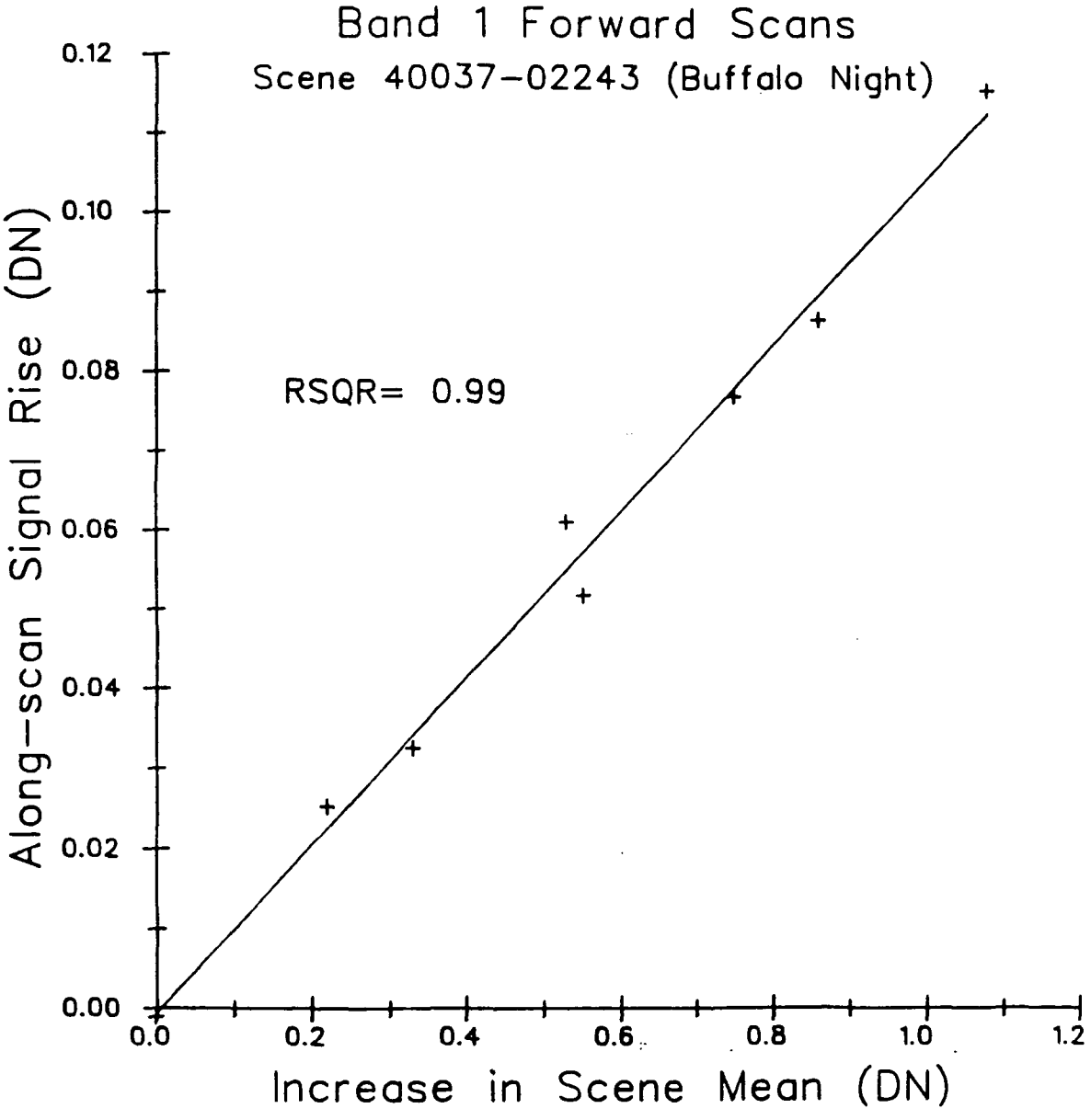


Figure 5. Relationship of Magnitude of Signal Rise and Difference Between Scan-cycle Mean and Scene Mean

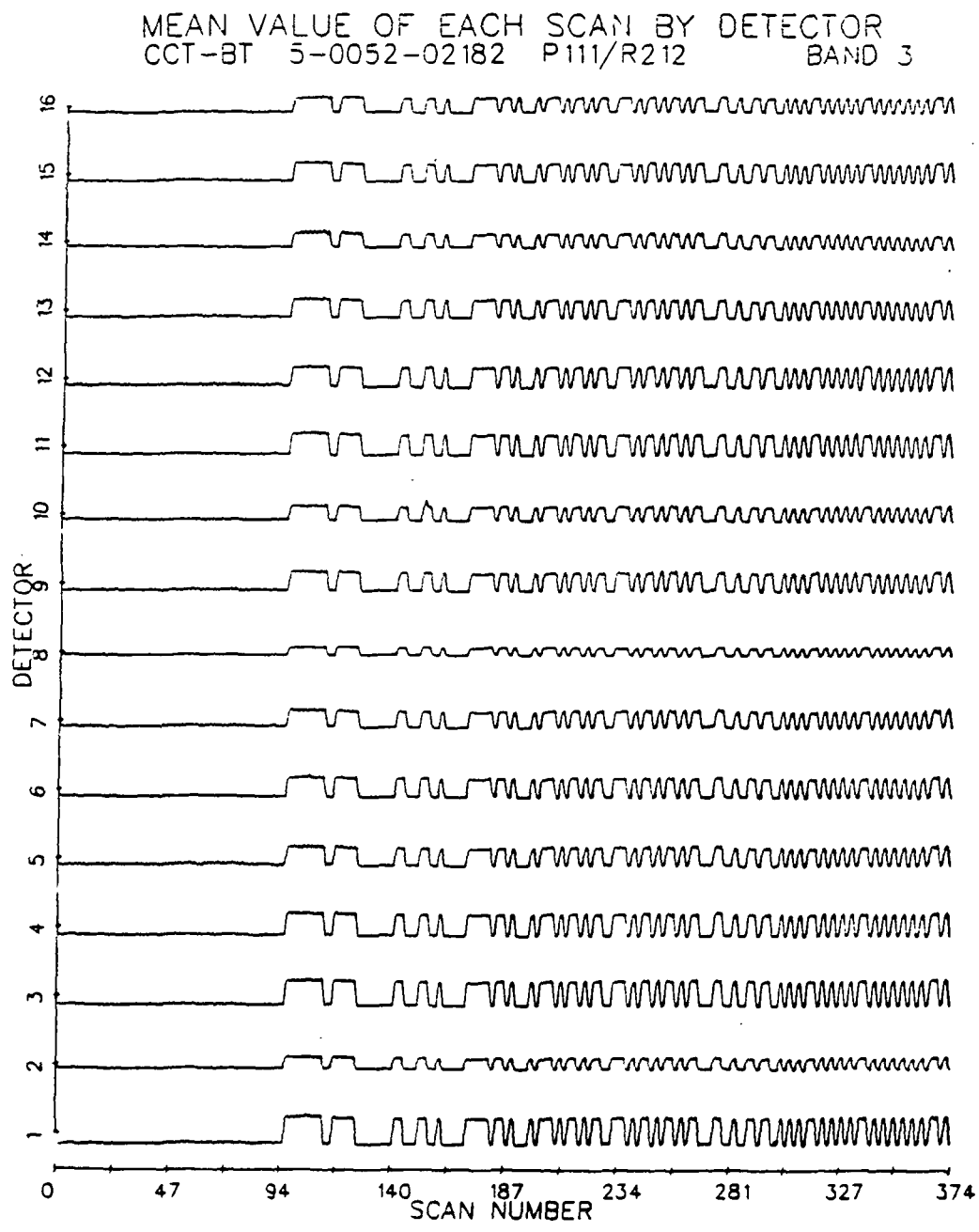
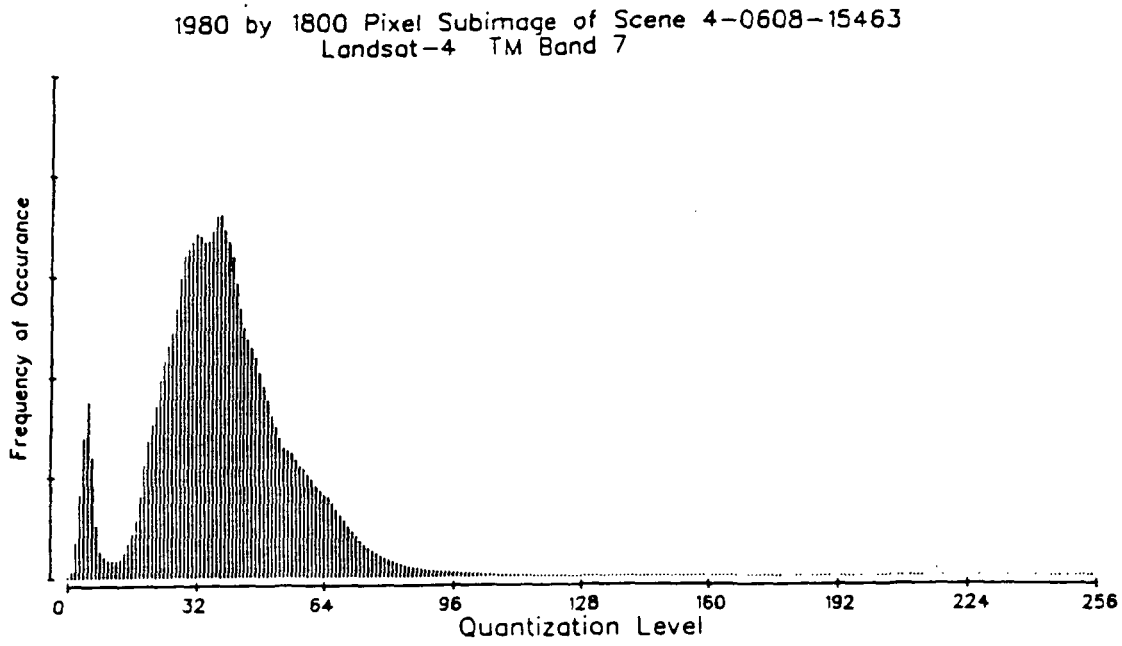


Figure 6. Level Shifts for Landsat-5 TM Band 3 Nighttime Data



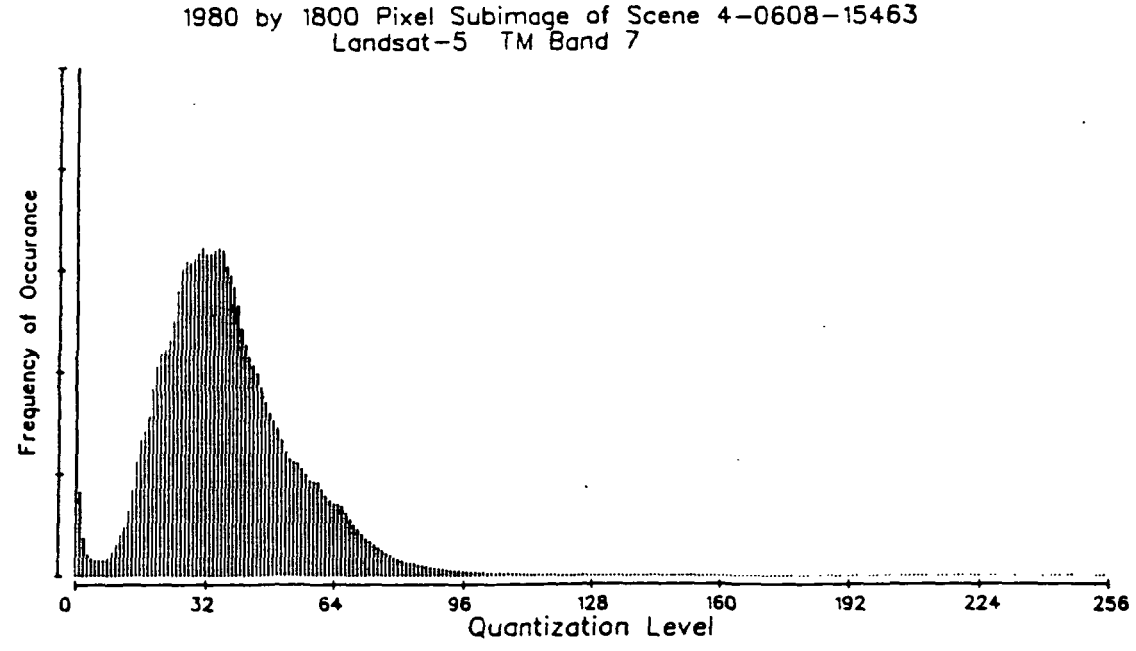


Figure 7. Landsat-4 and -5 TM Band 7 Histograms for Coincident Regions

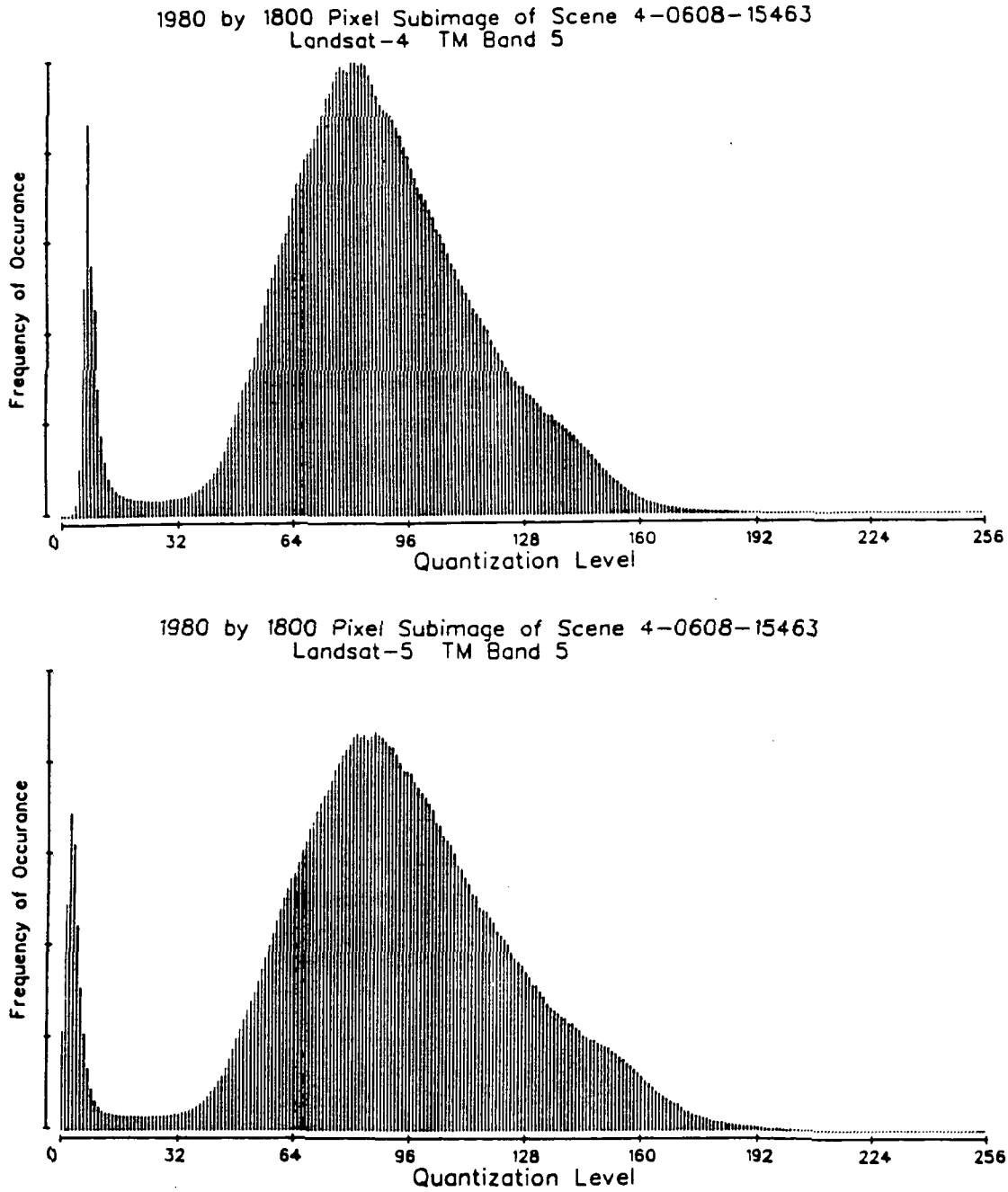


Figure 8. Landsat-4 and -5 TM Band 5 Histograms for Coincident Regions

APPENDIX B

Comparison of the Information Contents of Landsat TM and MSS Data[+]

Information-theoretic measures are applied to varied subsets of original and transformed data.

William A. Malila
Environmental Research Institute of Michigan
Ann Arbor, MI 48107

ABSTRACT

A communications-theory approach is taken to analyze the dispersion and concentration of signal values in various data spaces, irrespective of specific class membership. Entropy is used to quantify information, and mutual information is used to measure the information represented by subsets of spectral variables. Several different comparisons of information content are made. These include comparisons of system design capacities, of data volumes occupied by agricultural data in the spaces defined by original bands and by transformed spectral (Tasseled Cap) variables, of the information contents of original bands and Tasseled Cap variables, and of the information contents of TM and MSS for the given agricultural data sets. Also, the effects of sample size, scene content, and quantization level are examined.

[+]This research was sponsored by the U.S. National Aeronautics and Space Administration, Goddard Space Flight Center, Greenbelt, MD, under Contract NAS5-27346.

INTRODUCTION

In analyses of multispectral data sets produced by imaging remote sensing systems, needs arise for comparing the amounts of information provided by individual spectral bands, by various combinations of bands, and by different sensors. Measures based on classification performance or signal variance (e.g., principal component analysis) are commonly used for such comparisons. Classification procedures require knowledge of the identity of the scene elements being imaged and usually involve assumptions on the form of the signal distributions and parametric descriptors of those distributions. A class-independent and non-parametric measure of information content can be described in information-theoretic terms and is used here to analyze and compare digital image data from the Landsat Multispectral Scanner Subsystem (MSS) and Thematic Mapper (TM).

C. Shannon (1948) developed entropy measures of the information content of communications signals. Price (1984) and Bernstein, et al, (1984) made entropy calculations and comparisons of Landsat data on a band-by-band or component-by-component basis. Malila (1984) developed a procedure that takes into account dependencies among spectral bands and applied it to original and transformed versions of Landsat data; those results are summarized herein and extended to include additional data sets and other considerations.

METHOD

A communications-theory approach is taken to analyze the dispersion and concentration of signal values in various data spaces. Entropy, as defined by Shannon, is used to quantify information. The process of selecting a subset of bands is viewed as the transmission of data through a communication channel in which loss of information may occur, and the mutual information between input and output is used to measure information transfer, i.e., the information represented by the subset.

Several different comparisons of information content are made. These include (1) comparison of TM and MSS system-design information capacities, (2) comparisons of the TM and MSS data-space volumes spanned by the agricultural data in the spaces defined by both original bands and transformed spectral (Tasseled Cap) variables, (3) comparison of the agricultural information content of original bands to that of transformed variables, and (4) comparison of the agricultural information content of TM data to that of MSS. The effects of sample size and varied scene content are examined, as is the effect of coarser quantization.

BASIC INFORMATION CONCEPTS

Shannon defined self information, $I(x_i)$, as a measure of the information associated with knowing the occurrence of a signal state x_i which occurs with probability $P(x_i)$:

$$I(x_i) = \log_2\left(\frac{1}{P(x_i)}\right) = -\log_2 P(x_i) \quad \text{(bits)} \tag{1}$$

The more rare the event, the greater is one's uncertainty about when it will occur and, consequently, the greater is the information conveyed

when it is observed. Entropy, given the symbol H, is the value of self information when averaged over all N possible states of x:

$$H(x) = \sum_{i=1}^{N} P(x_i) \log_2 \frac{1}{P(x_i)} \tag{2}$$

Entropy is at its maximum when all states or cells are equally likely. It can be reduced by decreasing the number of cells occupied, by having a non-uniform distribution or a concentration of observations in the occupied cells, or both.

With two variables, the use of joint and conditional probabilities is necessary:

$$H(x,y) = H(x) + H(y|x) \tag{3}$$

since

$$P(x,y) = P(x)P(y|x) \tag{4}$$

In computing the conditional entropy, the weighting assigned to each information term is the joint probability of the states involved, i.e.,

$$H(x|y) = \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} P(x_i,y_j) \log_2 \frac{1}{P(x_i|y_j)} \tag{5}$$

If we consider x to be the input to a communication channel and y to be the output, we can define the mutual information transferred between them, i.e., $I_M(x;y)$, as

$$I_M(x;y) = H(x) - H(x|y) \tag{6}$$

This equation shows that the mutual information exchanged is the difference between H(x), the information content of the input, and H(x|y), the information loss or uncertainty about x when we are given the output y. When the

total information is transferred, $H(x|y) = 0$ and $I_M(x;y) = H(x)$. At the other extreme, when y does not contain any information relatable to x, $H(x|y) = H(x)$ and therefore $I_M(x;y) = 0$, i.e., there is no mutual information. Figure 1 presents a concise graphical summary of these quantities and their interrelationships. Numerical examples are given by Malila (1984).

MULTISPECTRAL EXTENSION

The above concepts can be extended to multispectral situations by letting the variables x and y become multidimensional vectors X and Y, with $X = (X_1, X_2, \ldots, X_{N_x})$ and $Y = (Y_1, Y_2, \ldots, Y_{N_y})$. Usually, $N_y \leq N_x$. The information transfer achieved by the communication channel is used here in a general sense, to represent both simple selections of spectral band subsets and more complex transformations, such as the Tasseled Cap Transformation.

ABSOLUTE VS. RELATIVE INFORMATION CONTENT

Multispectral sensors produce signals that have a fixed maximum number of signal levels in each spectral band, usually expressed as a number of bits, e.g., six bits for 64 levels in telemetered Landsat MSS bands and eight bits for 256 levels in Landsat TM bands. When the probabilities in the entropy equations are based on all possible combinations of those levels, absolute information measures will result. These would, for instance, be appropriate when absolute radiometric calibration of data is utilized.

Most current uses of multispectral data, however, employ techniques that utilize only relative amplitude information between signals from various scene elements. In these instances, the information resides in

the number of spectral cells that are occupied and the distribution of spectral values within them. Malila (1984) developed an expression that gives a relative entropy value, H_R, for any given data set, in terms of counts of occurrences of observations in cells of the spectral space. It is repeated here (for six variables):

$$H_R(X) = \underbrace{\log_2 N_{obs}}_{\substack{\text{Information if each observation were in a unique cell}}} - \underbrace{\left(\frac{1}{N_{obs}}\right) \sum_{ijklmn} \sum \ldots \sum C_{ijklmn} \log_2 C_{ijklmn}}_{\substack{\text{Information loss due to concentration of the observations into a subset of cells}}} \quad (7a)$$

where C_{ijklmn} is the count of occurrences in the cell having Level i in X_i, Level j in X_2, etc.,

and N_{obs} is the total number of observations in the data set being analyzed.

More briefly,

$$H_R(X) = H_{max} - H_{loss} \quad (7b)$$

It also is informative to divide the total information loss due to spectroradiometric concentration of signals (from Equation 7) into two components, one due to the reduced number of spectral cells which are occupied (below the total possible) and the remainder which occurs when the duplicate observations are not uniformly distributed among those cells, i.e.,

$$H_{loss} = L_{cell} + L_{unif} \quad (8)$$

where L_{cell} is the cell loss or loss in number of cells, i.e.,

$$L_{cell} = \log_2 N_{obs} - \log_2 N_{cells} = -\log_2 \left(\frac{N_{cells}}{N_{obs}}\right)$$

and L_{unif} is the uniformity loss,

$$L_{unif} = H_{loss} - L_{cell}$$

SPECTRAL BAND SUBSETTING

The selection of subsets of spectral bands is a special case of the mutual information expression,

$$I_M(X;Y) = H(X) - H(X|Y)$$

where Y now is a subset, X', of the X variables, so

$$I_M(X;X') = H(X) - H(X|X')$$

Whenever a variable, say X_p, is retained, its conditional probability term becomes unity, its contribution to H(X|X') is reduced to zero, and its information content is retained as mutual information. Whenever a variable, say X_q, is eliminated, there is a loss of mutual information. This loss is represented by the conditional entropy term through all conditional probability components in which X_q occurs on the left-hand side of the conditional probability indicator line but not on the righthand (or given) side.

SPECTRAL TRANSFORMS

Spectral transformations were obtained by applying the linear-combination Tasseled Cap (TASCAP) transformations to MSS [Kauth and Thomas, 1976] and six-band TM [Crist and Cicone, 1984] data. The principal TASCAP variables are Brightness and Greenness. The Brightness variables are positively weighted sums of all bands and respond to general changes in overall scene reflectance. The Greenness variables are essentially contrasts between near-infrared wavelengths (where healthy vegetation is more highly reflecting than soil) and visible wavelengths (where healthy vegetation tends to be less reflecting than many soils) and respond to the amount of vegetation present. These two variables capture 95 to 98% of

the variability in MSS data from typical U.S. agricultural scenes, while a third variable, called Wetness, has been found to be significant in similar TM data [Crist and Cicone, 1984]. The Tasseled-Cap variables, though related to principal-component variables, have advantages over them in that the Tasseled-Cap directions do not vary with the scene content and they have more consistent interpretability.

Also, principal-component analysis was utilized to obtain a different set of spectral variables for one comparison. All transformed values were rounded to the nearest integer before being analyzed.

QUANTIZATION EFFECTS

To explore the influence of quantization on the resultant information content, the amplitude values were re-quantized several times. At each step, the number of original digital counts per modified amplitude interval was doubled, thereby compressing the data and reducing the number of bits per channel by one for each step.

DATA SET

MSS and six-band TM data of two types were analyzed. These are (a) real Landsat-4 MSS and TM data acquired simultaneously from an agricultural scene in North Carolina and (b) data values synthesized from field-measured reflectance spectra of agricultural crops and soils using an atmospheric model. These data were used in prior comparisons of the spatial and spectral characteristics of Landsat TM and MSS data [Malila, et al, 1984 and Crist, 1984]. In the synthetic data, samples are primarily from vegetation at a variety of ground cover percentages, with many fewer examples of bare soil. All analyses of TM data are limited to the six

reflective bands; the thermal band was not analyzed in this effort due to its coarser spatial resolution, its dependence on emissive rather than reflective characteristics of scene materials, and lack of a comparable simulation data base. The TM frame was acquired on September 24, 1982, and included a wide range of agricultural crop conditions, ranging from bare soil to green and senescent vegetation to crop residues. It also included some samples from water and vegetation along the Atlantic Coast and from deciduous and coniferous trees.

RESULTS

SPECTRAL DATA VOLUMES

The diagram in Figure 2 helps describe the various terms used here to designate spectral data-space characteristics, while Table I quantifies many of the observed values. Figure 3 presents information measures for several of those quantities, as a function of the number of data variables. First, the system-design capacities of the Landsat-4 TM and MSS are presented, in terms of the number of bits transmitted to the ground and/or recorded on computer-compatible tapes (CCTs). For TM, the number of bits recorded on CCTs is the same as that transmitted (8 bits/channel). For MSS, however, the six-bit telemetered data are expanded to seven bits on the CCTs, with only an apparent gain of information. Nevertheless, many comparisons involving MSS will use seven-bit data since that is the form in which we received them. For some others, a degradation to six bits was performed before analysis. The greater information potential of the TM system design (reflective bands), as compared to the MSS system, is quantified as 48 vs. 24 bits in telemetered data.

Figure 3 also portrays the "hypercube" volume or data-space volume spanned by the TM and MSS data of Table IA. These volumes are computed by summing the bit equivalents of the observed data-value ranges (max - min + 1) in each band being considered. Upon comparing the fractions of their total data-space volumes that are spanned by data from the agricultural scene, one observes that the TM data fall nine bits short of capacity while the MSS data fall approximately six bits short of capacity.

Actual data dispersion volumes or relative entropies (see Figure 2 and Table I) were found to be substantially smaller than the hypercube volumes, due to correlations between bands and the limited numbers of observations. Results for the real TM data are shown in Figure 4 and for both TM and MSS (7 bits/band; CCT) in Figure 5. (Note that these relative-entropy values for actual information are substantially smaller than those reported by Price (1984) for similar comparisons in which the sum of band values was treated as the joint information content.) The data dispersion volumes in Figure 4 are measured by the relative entropies of the best variable combinations, and represent the relative information present in those sets. Most of the information is contained in the first two or three variables. Both the best and worst combinations are shown for each system in Figure 5. The number of observations analyzed establishes a maximum limit on each relative entropy value. As shown earlier in Equation (7), the concentration of multiple observations (pixels) into individual spectral cells reduces the information content below the potential maximum. Table I shows very little tendency for TM pixels to do this, due to the very large system capacity, spectral

diversity, and fine gradation of the TM bands. The MSS data show definite tendencies for multiple observations in spectral cells.

Table I shows that the TM data represent 3.3 bits more information than the MSS sensor data, with approximately two bits being associated with spatial resolution (pixel size and number) and the remainder with spectral bands and radiometric resolution. Since the synthetic data have the same number of observations for both TM and MSS, they can be considered to have equal spatial resolutions. Thus, the 2.2-bit difference must be solely due to their spectral and radiometric properties.

The above results were for a systematic sample from a larger area, 900 lines by 1300 TM pixels in size (450 x 650 MSS pixels). To explore the effects of sample size and scene content on the information measure, the area was divided into nine subareas, containing varied types and amounts of the scene classes. When all 1.17 million TM pixels were included in the analysis (Data Set TM-C), an information content equal to 18.4 bits of the possible 20.2 bits was computed, as shown in Table II. For the corresponding 0.29 million MSS pixels, 13.8 bits of the possible 18.2 bits were present as information. The two bits difference between maximum potentials is due to the greater number of TM pixels. Reductions below the maxima are due to reduced numbers of distinct spectral cells and non-uniformity of the cell populations. Bit equivalents of those losses also are indicated in Table II. It can be seen that substantially greater losses occur for MSS data than for TM data, leading to a total difference of 4.6 bits between the two data sets.

Values also were computed using all pixels in each subarea. Mean values are given in Table II (Data Sets TM-B and MSS-B), along with standard deviations to indicate the amount of variability found among the different scene areas. On the average, both types of losses are reduced from those found in the total data set, but variability among subareas is substantial. Even-smaller subsets of data were obtained for analysis by taking every tenth pixel in each subarea; the averages and standard deviations of those values also are listed in Table II (Data Sets TM-A and MSS-A). For these, the loss of information by TM is very minor (0.26 bit), but the losses for MSS remain greater (about one bit). Wharton (1984) simulated TM and MSS data sets, analyzed histograms of various sample sizes, and computed ratios of distinct to total number of samples. Comparable numbers were computed from the average cell loss values and are given in the last column of Table II as the percentage of cells which are distinct. The percentage for the largest real MSS data set is quite comparable to that for the largest set examined by Wharton, who found 27 percent distinct among 230,400 samples. His 59 percent for 28,800 samples and 85 percent for 3,600 samples are higher than the respective 34 percent in Table II for 32,500 samples and 66 percent for 3,300 samples. For TM, Wharton found nearly 100 percent distinct cases for even the 230,400 sample case, versus 61 percent here for 130,000 samples, but he considered seven rather than the six dimensions analyzed here and included only samples from nine scene classes. The TM data in Table II retain much more distinctness than MSS as the sample size is increased, with 89 percent distinct for 13,000 samples, 61 percent for 130,000 samples, and 58 percent distinct for 1,170,000 samples.

SPECTRAL TRANSFORMATIONS

Figure 4 also compares the data-space volumes spanned by original bands and Tasseled-Cap transformed versions of signals from the agricultural scene (Table IA sample). Three fewer bits per pixel are required to provide the same information using the transformed variables than would be required by the original bands. This effect potentially could be used to reduce telemetry requirements; differences might be even greater for data sets with a broader range of scene amplitudes.

For the synthetic MSS data set, a comparison was made of the information content of original band values and two types of transformed variables, TASCAP variables and principal-component variables. They were found to be essentially identical. The equality of the complete sets of variables is in keeping with theoretical considerations of linear transformations.

To compare with the original-band values of Figure 5, relative entropy values for the best and worst TASCAP subsets of each size are presented in Figure 6. In this case, we find an even greater disparity between best and worst combinations, due to the decreased information content of the last TASCAP variables. Here again, relatively little information is gained by the inclusion of more than three variables.

DIMENSIONALITY

Figure 7 displays relative entropy values computed for the first three Tasseled-Cap components of TM and MSS data from the agricultural scene (Table IA sample). (The MSS data were in CCT form at seven bits/band.) The first three components are individually quite similar for TM, but

there is a substantial decrease (3.3 bits below Brightness) for the third component of MSS (Yellowness). This is consistent both with many investigators' experiences in finding MSS data of agricultural areas to be primarily two dimensional and with recent studies which have found a substantial amount of information in the TM Tasseled Cap Third Component [Crist and Cicone, 1984]. Throughout this comparison, TM values are greater than the corresponding MSS values, for example the TM Brightness value is 6.7 bits compared to 5.8 bits for MSS.

When pairs of components are considered, we see substantial increases in total information, as would be expected with the addition of a second variable; the value for TM Brightness/Greenness is 4.8 bits greater than for Brightness alone, and the corresponding increase for MSS is 3.7 bits. However, differences do appear between MSS and TM. Whereas the value of the Brightness/Greenness pair for MSS is substantially greater than the other two (approximately two bits greater than Greenness/Third Component), there is relatively little difference (less than 0.4 bits) among the three pairings from TM data, pointing to a higher dimensionality in TM.

Three components captured the vast majority of information for both systems. However, the fact that the gain in going from two to three components was nearly as large for MSS (1.25 bits) as for TM (1.70 bits) was somewhat surprising in view of the previously discussed two-dimensional character of MSS data. Furthermore, principal-component analysis of MSS data showed nearly total representation of variance by the first two components. The MSS gain likely is due to the Brightness/Greenness plane having a thickness of several counts in the third direction, even though this third component was uncorrelated with the others. The observed

values also indicate that differences do exist among these various measures of multispectral signal properties. The TM data pattern also may be somewhat planar in three space, although not aligned as well with any component axis; correlations with the Third Component were -0.69 for Brightness and 0.36 for Greenness in this data set. None of these observations, however, should diminish the utility of Tasseled Cap transforms for physical interpretation of data values and agricultural scene characteristics.

NOISE

Noise in multispectral data was not considered explicitly in the results presented thus far. Sensor noise effects certainly were present in the real Landsat data and natural variations of crop observations were present in both the real and synthetic data. Noise can add variance to signals and increase the number of spectral cells occupied (above that for no noise), thereby creating an apparent information content greater than the true information content of ideal, noiseless signals. To explore such effects, the number of discrete levels present in the data sets was reduced by applying several different quantization factors (greater than unity) to each band and computing the reduced information content. The results are summarized in Table III for three subareas which had (relatively) high, medium, and low information content, respectively. The TM still had more information when degraded to seven bits per band but, by the time the amplitude data were degraded to five bits per band, there was little difference between the corresponding TM and MSS data sets.

SUMMARY

An information-theoretic measure was defined and used to compare Landsat MSS and TM multispectral data. The measure quantifies signal dispersion patterns, independently of class membership and distributional assumptions. It provides an alternate method (to classification) of measuring the extent to which subsets of bands or transformed variables represent the total pattern. The relative entropy value is limited by the number of observations being analyzed. Since results do vary with scene content, analysts should insure that data sets being analyzed are representative of the problems under consideration.

A number of observations were made. The TM system-design information capacity is much greater than that of MSS. The potential information capacities and the signal "hypercube" volumes of agricultural data were much larger than the information actually represented by signal dispersion patterns in the sets of data values analyzed. Tasseled Cap transformations preserved the information in original bands and offered a modest savings in bits over those original bands, a fact which might be useful in data compression approaches. Relatively few multiple occurrences of spectral observations were found in the TM data sets compared to MSS, another indication of TM's finer partitioning of spectral space. For the "best" combinations of variables, relative entropy magnitudes were more a function of the number of variables than of the type of variables (original bands or transformed). TM had greater relative entropy values for Brightness and Brightness/Greenness than did MSS. Information in the Tasseled Cap Third Component of TM was much greater than that of MSS,

both by itself and in combination with Brightness or Greenness, confirming TM's greater dimensionality. Reductions in the number of bits used to encode data in each channel decreased the information content, affecting TM data proportionately more than MSS data so that, with five bits or less per band, the information in comparable sets was equal.

REFERENCES

Bernstein, R., J. Lotspiech, H. Myers, H. Kolsky, and R. Lees, 1984. "Landsat-4 MSS and Thematic Mapper Data Quality and Information Content Analysis", IEEE Transactions on Geoscience and Remote Sensing, pp. 222-236.

Crist, E.P., 1984. "Comparison of Coincident Landsat-4 MSS and TM Data Over an Agricultural Region", Technical Papers of the 50th Annual Meeting of the American Society of Photogrammetry, Vol. 2, pp. 508-517.

Crist, E.P. and R.C. Cicone, 1984. "A Physically-based Transformation of Thematic Mapper Data - The TM Tasseled Cap", IEEE Transactions on Geoscience and Remote Sensing, pp. 256-263.

Kauth, R.J. and G.S. Thomas, 1976. "The Tasseled-Cap -- A Graphic Description of the Spectral-Temporal Development of Agricultural Crops as Seen by Landsat", Symposium on Machine Processing of Remotely Sensed Data, June 29 - July 1, 1976, Purdue University, LARS, W. Lafayette, IN.

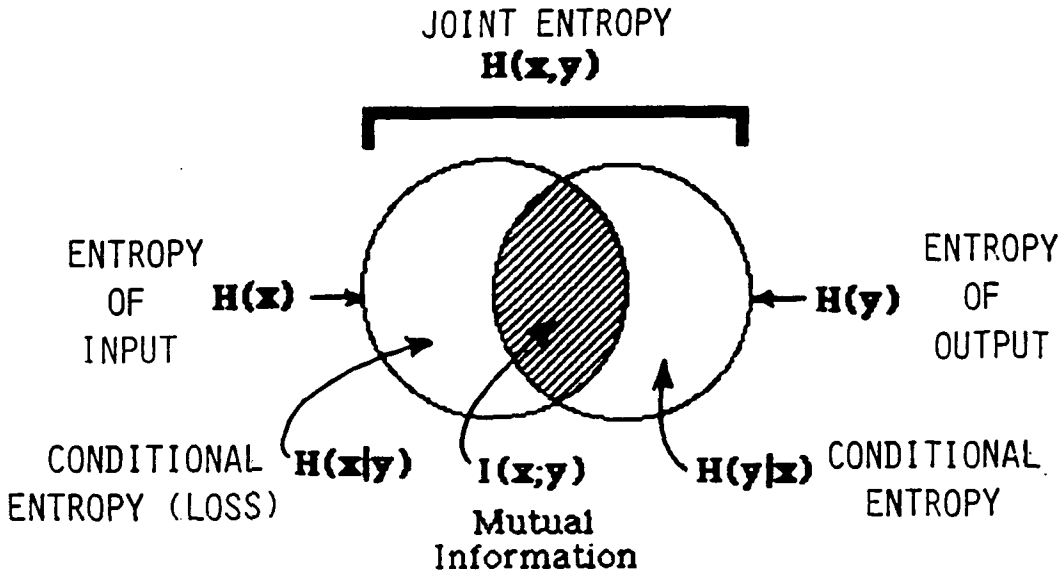
Malila, W.A., M.D. Metzler, D.P. Rice and E.P. Crist, 1984. "Characterization of Landsat-4 MSS and TM Digital Image Data", IEEE Transactions on Geoscience and Remote Sensing, pp. 177-191.

Price, J.C., 1984. "Comparison of the Information Content of Data from the Landsat-4 Thematic Mapper and Multispectral Scanner", IEEE Transactions on Geoscience and Remote Sensing, pp. 272-281.

Wharton, S.W., 1984. "An Analysis of the Effects of Sample Size on Classification Performance of a Histogram-based Cluster Analysis Procedure", Pattern Recognition, Vol. 17, No. 2, pp. 239-244.

FIGURES

TABLES



Summary of Information Relationships

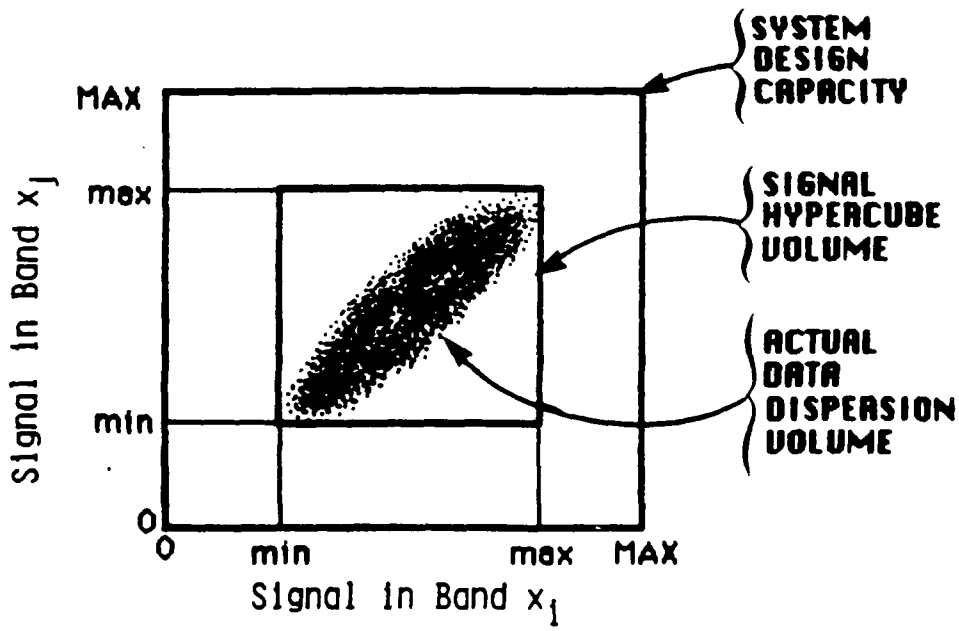
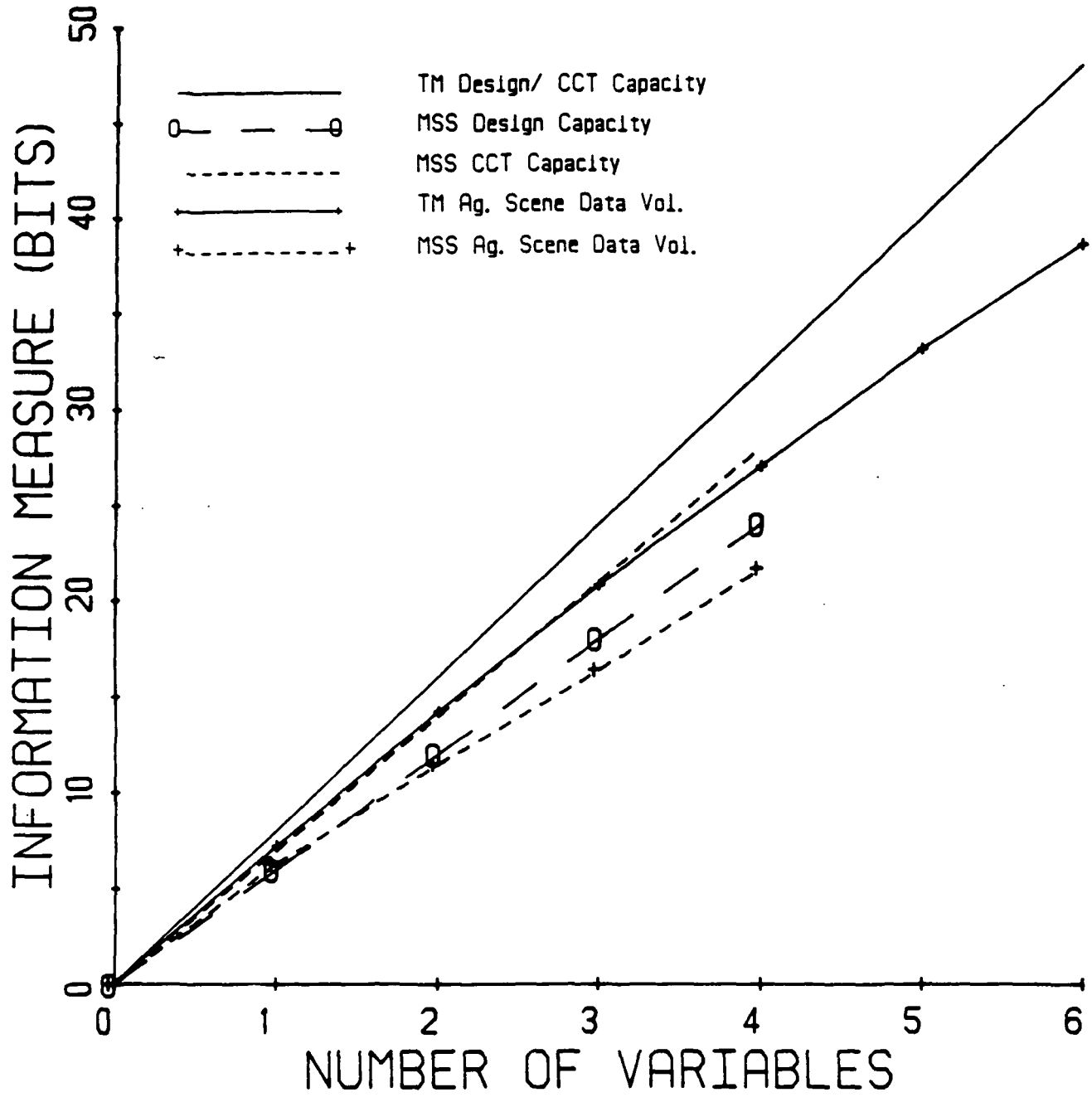
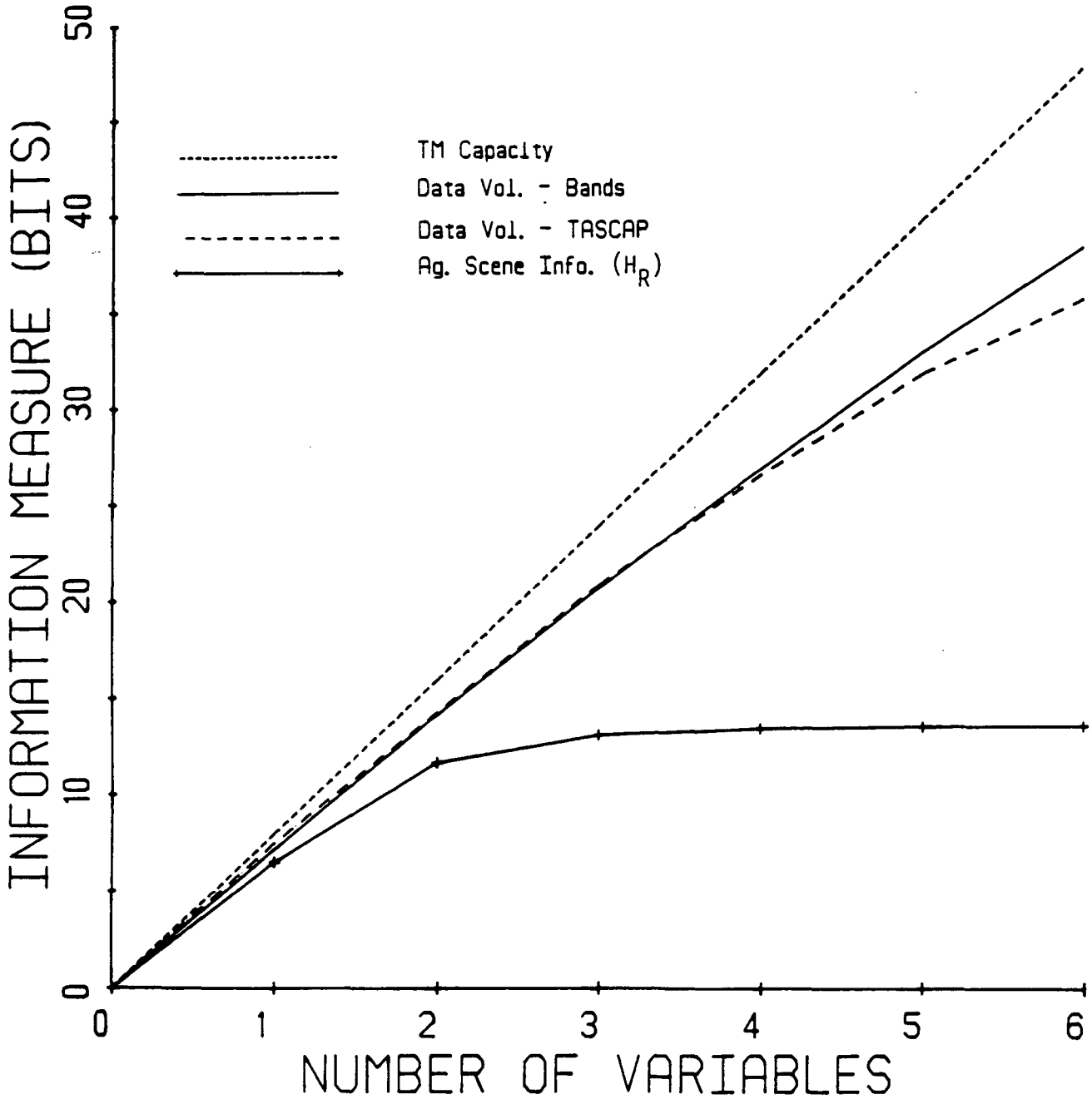


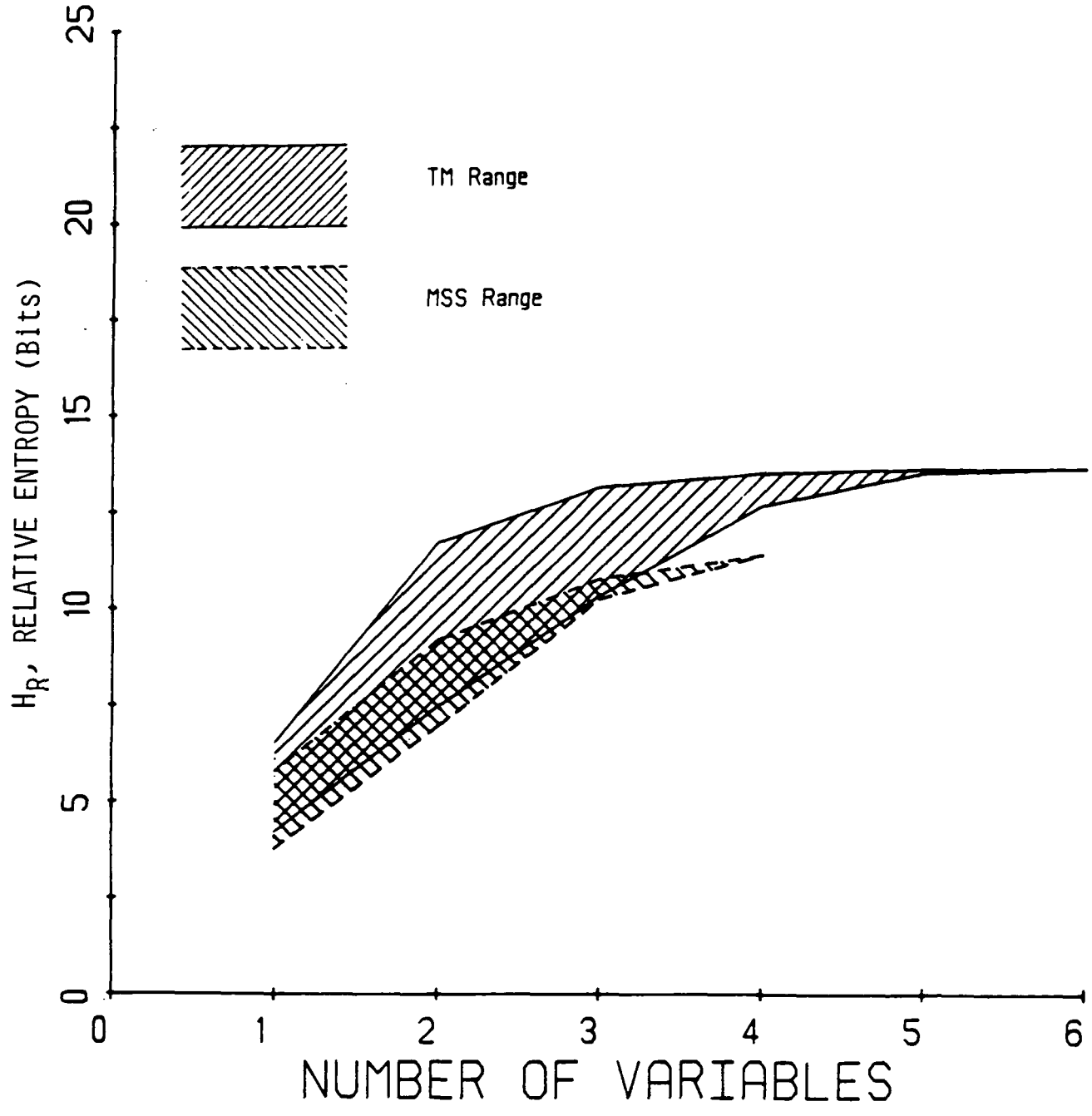
Illustration of Various Spectral Data Volumes



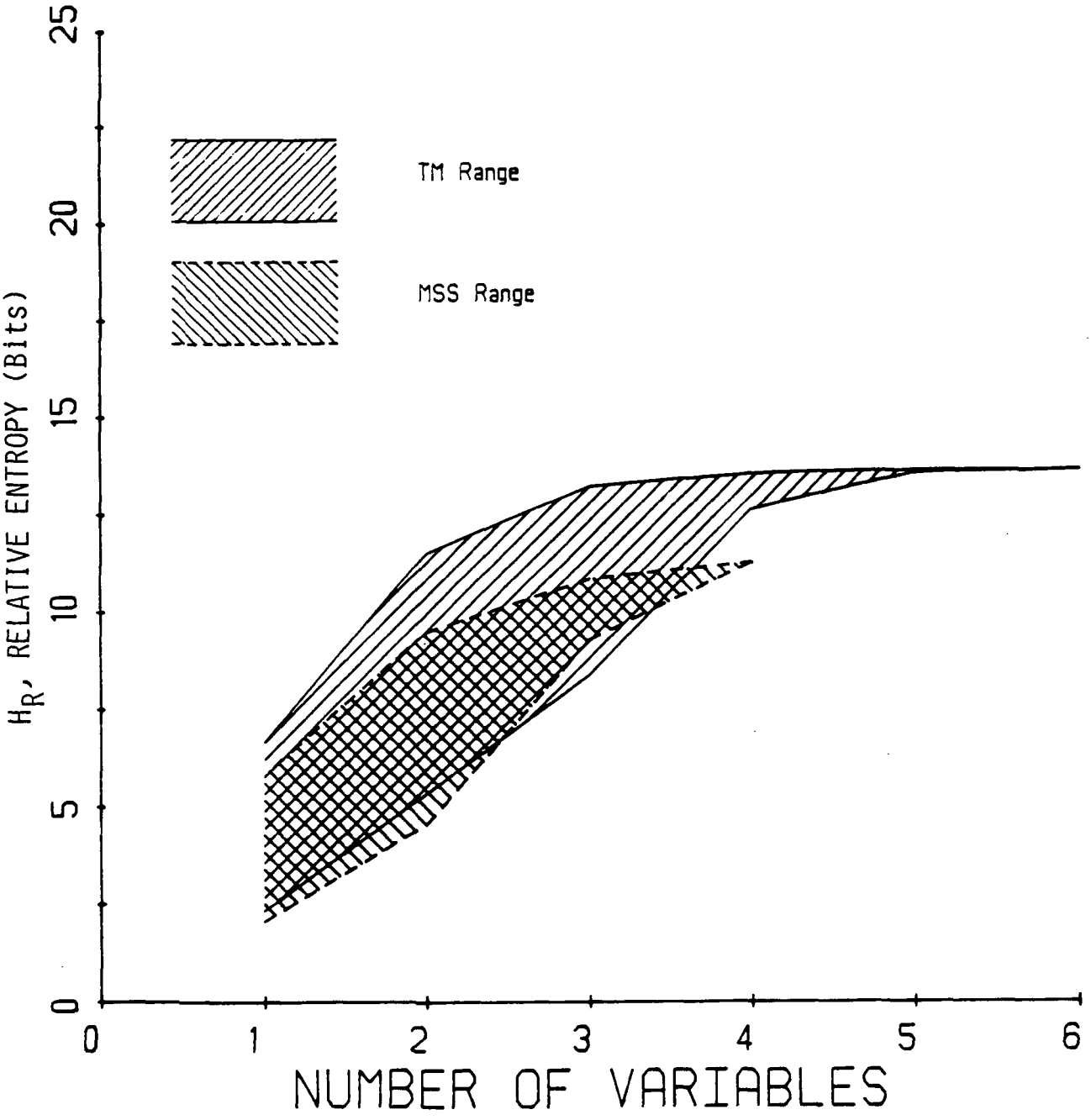
Comparison of Landsat TM and MSS Information Capacities



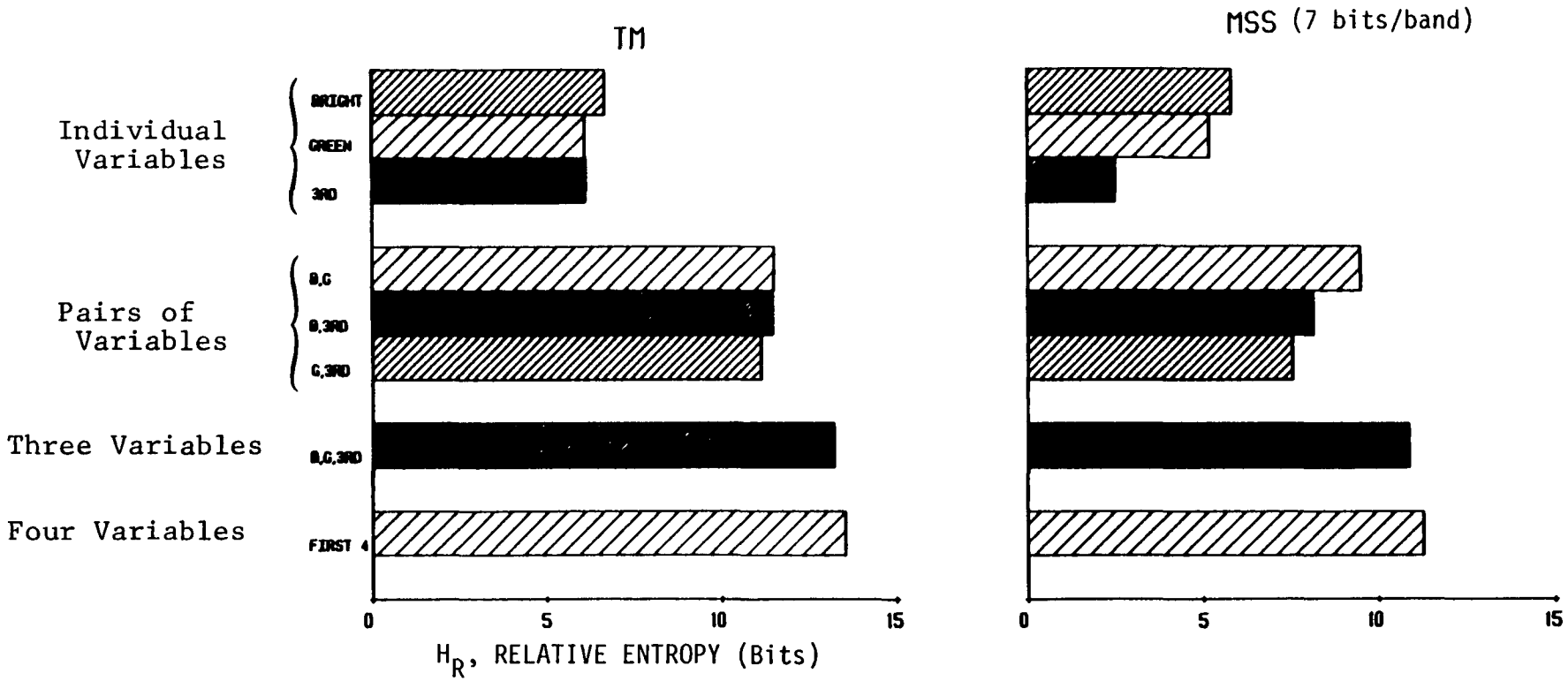
Thematic Mapper's Utilization of Data Space



Range of Information in Subsets of Bands



Range of Information in Subsets of Tasseled-Cap Variables



Comparison of Information Contents of TM and MSS Tasseled-Cap Variables

Information Comparison for MSS and Six-Band TM Data Sets

A. VALUES FOR REAL AGRICULTURAL DATA

(A common area from the N. Carolina scene)

		MSS		SIX-BAND TM		TM GAIN
		NUMBER	BITS	NUMBER	BITS	(BITS)
SYSTEM CAPACITY:	SENSOR	0.17×10^8	24	0.28×10^{15}	48	24
	CCT	0.27×10^9	28	0.28×10^{15}	48	20
HYPERCUBE VOL.:	SENSOR	0.44×10^6	18.7	0.43×10^{12}	38.6	20
	CCT	0.32×10^7	21.6	0.43×10^{12}	38.6	17
DATA DISPERSION PATTERN:						
	· # Observations; H_{max}	3,468	11.8	13,015	13.7	1.91 (Spatial)
MSS CCT: 7 Bits per Band	· # Unique cells	2,898	-	12,903	-	
	· Relative Entropy, H_R	-	11.4	-	13.7	2.27 (Total)
	· Entropy loss due to spectral concentration, H_{loss}	-	0.38	-	0.02	0.36 (Spectral)
MSS Sensor: 6 Bits per Band	· # Unique cells	1,730	-	12,903	-	
	· Relative Entropy, H_R	-	10.3	-	13.7	3.34 (Total)
	· Entropy loss due to spectral concentration, H_{loss}	-	1.45	-	0.02	1.43 (Spectral)

B. VALUES FOR SYNTHETIC AGRICULTURAL DATA

(Assumes equal spatial resolution)

	MSS		SIX-BAND TM		TM GAIN
	NUMBER	BITS	NUMBER	BITS	(BITS)
"SYSTEM" CAPACITY (MSS: 6 Bits/Band)	0.17×10^8	24	0.28×10^{15}	48	24
OBSERVED HYPERCUBE VOLUME	0.10×10^7	20	0.99×10^{12}	40	20
DATA DISPERSION PATTERN:					
· # Observations; H_{max}	2,276	11.15	2,276	11.5	
· # Unique cells	817	-	2,260	-	
· Relative Entropy, H_R	-	8.94	-	11.14	2.20 (Spectral)*
· Entropy loss due to spectral concentration, H_{loss}	-	2.21	-	0.01	

*(TM gain over seven-bit simulated MSS data was one bit.)

Effects of Sample Size and Scene Diversity on Information Content

Data Set	Number of Pixels	H_{max} Maximum Possible Relative Entropy (bits)	H_R Actual Relative Entropy (bits)	L_{cell} Loss in Number of Cells (bits)	L_{unif} Uniformity Loss (bits)	$\frac{N_{cells}}{N_{obsv}} \times 100$ Percent Distinct Cells+
TM-A	1.3×10^4	13.67	13.41*	0.162*	0.091*	89.4
			(0.21)	(0.136)	(0.078)	[75 - 99]
TM-B	1.3×10^5	16.99	15.66*	0.711*	0.615*	61.1
			(1.37)	(0.675)	(0.704)	[38 - 96]
TM-C	1.17×10^6	20.16	18.41	0.791	0.954	57.8
MSS-A	3.3×10^3	11.69	10.72*	0.604*	0.361*	65.8
			(0.47)	(0.331)	(0.148)	[44 - 88]
MSS-B	3.25×10^4	14.99	12.27*	1.539*	1.179*	34.4
			(1.04)	(0.693)	(0.380)	[16 - 61]
MSS-C	2.93×10^5	18.16	13.81	2.149	2.200	22.5

* Denotes mean of values from nine subareas.

( ) Denotes standard deviation of those values.

\+ Computable from average bits of cell loss, i.e., $100 \times 2 \exp(-L_{cell})$.

[ ] Denotes range of values computed for individual samples.

Effects of Quantization Detail on Information Content

Sensor	Number of Pixels	Relative Scene Complexity	Relative Entropy (bits) for Indicated Number of Bits per Band: 8	7	6	5	4	3
TM	1.3×10^5	High	16.9	15.6	12.3	9.0	6.2	4.6
	"	Medium	16.1	13.4	10.0	7.2	4.8	3.4
	"	Low	15.3	11.8	8.5	5.8	3.7	2.7
MSS	3.3×10^4	High	-	13.8	10.6	9.1	6.4	4.3
	"	Medium	-	12.0	9.9	7.2	4.9	3.3
	"	Low	-	11.0	8.7	6.1	3.8	2.5

Sensor	Relative Scene Complexity	Fraction of Maximum Relative Entropy for Indicated Bits/Band: 8	7	6	5	4	3
TM	High	1.00	0.93	0.73	0.53	0.37	0.27
	Medium	1.00	0.83	0.62	0.44	0.30	0.21
	Low	1.00	0.77	0.56	0.38	0.24	0.17
MSS	High	-	1.00	0.76	0.66	0.46	0.31
	Medium	-	1.00	0.83	0.60	0.41	0.28
	Low	-	1.00	0.79	0.55	0.35	0.22

DISTRIBUTION LIST

NASA Goddard Space Flight Center
Greenbelt Road
Greenbelt, Maryland 20771

	Copies
Contracting Officer, Code 284.4	1
Publication Branch, Code 253.1	1
Patent Counsel, Code 204	1
Technical Officer, Mr. Harold Oseroff, Code 902	10